Dispersion and current measurements. An investigation based on time series analysis and turbulence models

Boelskifte, S.

Publication date:
1986

Document Version
Publisher's PDF, also known as Version of record

RISØ-M-2566

DISPERSION AND CURRENT MEASUREMENTS

An investigation based on time series analysis and turbulence models.

Søren Boelskifte

Abstract. A model for the simulation of particle movements in water should incorporate the mutual distance dependent correlation. As long as reliable data are accessible a model can be created of the dispersion in a given area from a statistical description of turbulence. Current measurements have been performed in an area north of the Swedish nuclear power plant Barsebäck, and statistical time series analysis have made it possible to estimate multivariate autoregressive moving-average (ARMA) models for these data using the Box-Jenkins method. The correlation structure for the area has been investigated in detail.

Transport and dispersion models for the marine environment are used in estimating doses to the population from the aquatic food chain. Some of these models are described with special emphasis on the time and length scales they cover.

Furthermore, to illustrate the background of the simulation model, short introductions are given to health physics, time series analysis, and turbulence theory.

(Continued next page)

April 1986
Risø National Laboratory, DK-4000 Roskilde, Denmark

Analysis of the simulation model shows the relative importance of the different parameters. The model can be expanded to conditional simulation, where the current measurements are used directly to simulate the movement of one of the particles.

Results from the model are also compared to results from a sampling of bioindicators (Fucus vesiculosus) along the Danish coast. The reliability of bioindicators in this kind of experiment is discussed.

The present thesis is submitted to the Technical University of Denmark in partial fulfilment of the requirements for the degree of Lic.Techn. (Ph.D.). Associate Professor Henrik Spliid from the Institute of Mathematical Statistics and Operations Research (IMSOR) acted as responsible supervisor. Dr.Scient. Asker Aarkrog functioned as supervisor.

Subtle is the Lord,
but malicious He is not.
Einstein.

ISBN 87-550-1216-7
ISSN 0418-6435

Grafisk Service 1986

CONTENTS

Page

1. INTRODUCTION

Health physics occupies a unique and important place among the environmental sciences. Concerned primarily with radiation protection, it encompasses such topics as dose calculations and activity measurements. In contrast to other types of pollution concentration, radioisotopes can be detected at extremely low levels. Since there exists a well established relationship between a received radiation dose and the resultant detriment, it is possible to determine the health effect of any amount of activity.

Because they can be detected at such low levels, radioisotopes can be used to trace the transport and dispersion of other forms of pollution, and, in some cases, to test different transport models.

Detailed investigations have been made concerning the transport and dilution of radioactive materials in the atmosphere and the terrestrial environment.

Interest in aquatic pathways has been less pronounced; probably the risk related to water-borne releases is less acute than that related to atmospheric releases, since transport is much faster and less controlled in the air.

The form of a model for water transport is sensitive to the time and length scale considered, but in any case the dilution mechanics should be evaluated. Thus, the concept turbulence is an important factor to analyse in the model construction. Turbulence is stochastic in its nature; the analysis therefore should be based on statistical terms.

In general, models should be handled with great care, with separate analysis of each model's assumptions. It is important to know whether a model is based on theoretical explanation, on measurements, or both. The possibility of verification is also

relevant. Depending on their purpose, some models are easy to verify experimentally. Whereas some can be impossible to verify, either because they cover too long a time span, or because they result in an unlikely occurrence (a "100 year wave", a reactor accident, etc.). The exactness of results, which can differ considerably among models, should also be taken into account.

Time series analysis, a statistical tool which has been used in fields ranging from meteorology to economics to medicine, can be useful in describing the development of stochastic processes over time. In this report a time series model for turbulence is constructed based on current measurements performed especially for the purpose of providing data for that model.

The structure of the model is based on a theoretical description of turbulence. The parameter values are based on the current measuring program. Our aim in producing this model is to simulate the dilution of a discharge of radioactive material into the marine environment. Operation of this model will produce various statistics, including mean and maximum concentrations. With the use of bioindicators some aspects of the model, such as long-term average dilution, can be verified. Thus, the objects of the study have been to develop methods for model building using time series analysis and to test the applicability of such models in practical health physics.

As can be seen above, the report contains elements of turbulence theory, time series analysis, and health physics. A short introduction to each area is therefore given (Chapter 2). The selection of marine environment transport models found in Chapter 3, illustrates diverse purposes and methods, including the demand for data. The current measurements, which provide the basis for the simulation model, are described in Chapter 4, where the important correlation analysis is also performed. In Chapter 5 we explain how time series analysis (Box-Jenkins method) can be used to describe a current. The results are used in Chapter 6, and the simulation model is specified. A verification of some aspects of the model is found in Chapter 7, where the reliability of bioindicators is also examined.

2. THEORETICAL BACKGROUND

2.1. Turbulence and diffusion

2.1.1. Introduction

In the chapters that follow, some elements of the theory of turbulence and diffusion are applied, and will therefore be mentioned briefly here. Several textbooks on the topic exist and for further studies reference can be made to Fischer et. al (1979) and Lumley and Panofsky (1964), among others. Rather than providing a detailed introduction, this section serves merely as a reference for the rest of the report.

Most authors point out that a precise and common definition of turbulence cannot be given. Instead, turbulence is described by some of its properties. Usually the current (or wind) velocity is expressed in the following way:

$$u_t = U + u' \tag{2.1}$$

where u_t is the current velocity, U the mean velocity, and u' the turbulent, irregular fluctuations around this mean. In contrast to laminar flow, turbulent flow is characterised by a large Reynolds number $R = UL/\nu$, where U is the mean velocity, L a characteristic length scale, and ν the viscosity.

Turbulence is stochastic in its nature, which means that predictions and descriptions of the flow must be based on statistical terms. The equations of motion have not been shown to yield a unique solution for a given set of conditions, and in practice, no matter how carefully the conditions of an experiment are reproduced, the velocity field cannot be predicted in detail.

Another important feature of turbulence is its ability to disperse properties. In laminar flow, a dye thread will exist for a long time; whereas in turbulent flow, it will soon grow indis-

tinct. This phenomenon partially explains the wide-spread interest in describing turbulence, and is also the topic of the present study.

A discharge into the sea is affected by a current in two ways according the aforementioned equation (2.1), where U represents the important average transport of the center of mass, and u' is responsible for the diffusion around this center. Often the word dispersion expresses advection plus diffusion, where advection means "transport by an imposed current system, as in a river or coastal waters" (Fischer et. al. 1979). Sometimes it is used to indicate the effect of both turbulent diffusion and shear, i.e. the diffusing effect of a velocity profile.

In Chapter 4 we discuss the question of separating U and u'.

2.1.2. Turbulence

In creating a statistical description of turbulence, we must consider the concept of correlation. The cross-correlation between the velocities of two points, P and P', is divided into a longitudinal and a transverse component. If r is the distance between the two points, u_1 and u_1' the velocities in the direction of the line between the two points, and u_2 and u_2' velocities perpendicular to this line (Fig. 2.1), the longitudinal correlation f(r) is defined as

$$f = f(r) = \frac{E\{u_1 u_1'\}}{\underline{u}_1^2} \tag{2.2}$$

where $\underline{u}_1^2$ is the variance of u_1.

Similarly, the transverse correlation g(r) is defined as

$$g = g(r) = \frac{E\{u_2 u_2'\}}{\underline{u}_2^2} \tag{2.3}$$

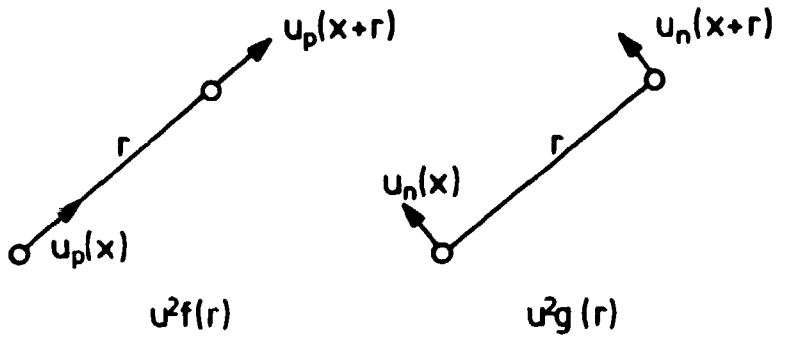


Fig. 2.1. A symbolic representation of the longitudinal and transverse correlations (Lumley and Panofsky, 1964).

For both correlations it is intuitively clear that the value tends towards 1 as r approaches 0 and tends towards 0 as r approaches infinity. In case of isotropy a relation exists between f and g:

$$g(r) = \frac{1}{2r}\frac{\partial}{\partial r}(r^2 f) \qquad (2.4)$$

Assuming an exponential function for f, $f(r) = \exp(-r/L)$ this relation leads to $g(r) = (1-r/2L)\exp(-r/L)$.

One problem of major concern in dispersion invesitigations is the difference between Eulerian and Lagrangian description. Eulerian current measurements, i.e. measurements from a single point, are easy to obtain, whereas Lagrangian measurements, i.e. measurements related to a single particle following the current, are much more difficult to obtain. But the problem of transport and diffusion is most easily interpreted in Lagrangian terms; however, it is connected with particles, rather than points. Many authors have been interested in the theoretical problem of relating the two descriptions, e.g. Kofoed Hansen and Wandel (1962). The Eulerian autocorrelation f_E is defined as

$$f_E(\tau) \cdot \underline{u}^2 = \overline{u_1(t)u_1(t+\tau)} \qquad (2.5)$$

where the overbar indicates average value. The Lagrangian autocorrelation f_L is expressed in the same way, but u is now related to a fixed particle. Hay and Pasquill (1959) give the hypothesis

$$f_E(\tau) = f_L(\beta\tau) \qquad (2.6)$$

which is intuitively understandable since it indicates that the only difference between the two formulas is a matter of scale. Changes in the velocity of a particle takes place more slowly than changes at a fixed point.

Many expressions for β exist, e.g. $\beta = \underline{u}_1/\alpha u_i$ (Engelund,1969) and $\beta = \sqrt{\pi}/4\ i^{-1}$, (Kofoed Hansen and Wandel, 1962) where i is the intensity of the turbulence defined as $\underline{u}/U$.

To simplify the description of the current field some assumptions usually are made. The turbulence often is said to be stationary, i.e. the mean and all moments of higher order are constant in time. Also homogeneity is usually assumed, i.e. the statistical properties of the flow are the same at all spatial points in the area considered. These two assumptions imply that the variance of the velocity is steady and does not change with time or position. The third frequently used assumption is isotropy, i.e. the properties of the turbulence are unchanged by a rotation of the coordinate system.

Lumley and Panofsky (1964), among others, have an introduction to the spectral theory of turbulence. One important result of applying the spectrum is the "Kolmogorov law" or the "-5/3 law". It states that in the inertial subrange, where no production or dissipation takes place, i. e. there is only inertial transfer to smaller and smaller eddies, the spectrum must have the form

$$E(n) = \alpha\ \varepsilon^{2/3}\ n^{-5/3} \qquad (2.7)$$

where n is the frequency, α a constant and ε the dissipation of turbulent velocity fluctuations by molecular friction. An example of such a spectrum is given in Fig. 4.11.

Usually meteorologists describe their spectra by means of a logarithmic frequency scale and an ordinate which represents the product of frequency and spectrum estimate nS(n). Thus, the area between two frequencies expresses variance contributed by the frequency interval.

2.1.3. Diffusion

After this very short introduction to some elements of turbulence theory, its application to diffusion problems will be described. The discussion here is mainly based on Fisher et. al. (1979), but many other references have been used (Mikkelsen, (1982), Gryning (1981), Odgaard (1975), and others).

Subsequently, we will provide a more stringent explanation of the ambiguous term "Spread of pollution." First of all we shall consider a large number of particles released at the same time and placed in a turbulent flow. For convenience, only two dimensions will be considered and the mean current will be ignored by choosing a coordinate system following the mean current. After some time, the particles will form a "cluster" of uneven shape, with a new center of mass (see Fig. 2.2). A repeated trial would result in another cluster shape and yet another center of mass.

Two interpretations of repeated trials exist. One approach is to average over all of the releases to obtain an ensemble average. At a fixed point in space this may result in a number of zeros plus a few large values. "This type of average may be meaningless to an organism subject to release of a pollutant, because the few high concentrations may kill the organism and the large number of zeros cannot bring it back to life" (Fischer et. al. 1979). On the other hand, assuming non-fatal concentrations, seaweed for instance presumably performs this integration with at least some radionuclides, (described in detail in Chapter 7). Another ap-

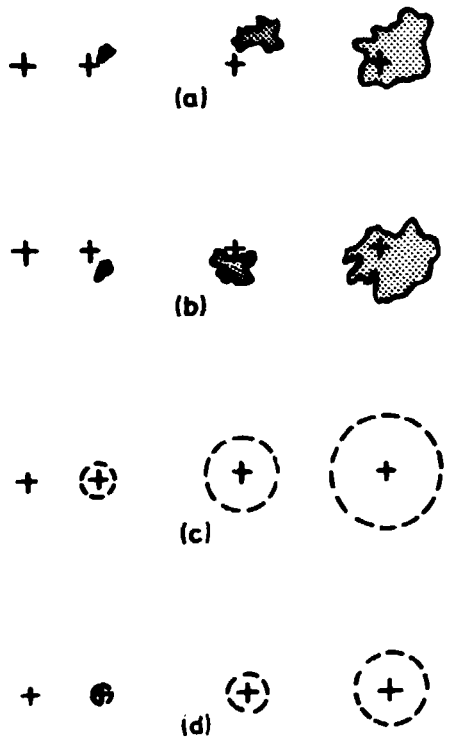


Fig. 2.2. Diffusion in homogeneous, isotropic turbulence with zero mean velocity. The largest scale of motion is slightly larger than the largest clouds shown. (a) Spread of a single cloud. (b) Spread of a second cloud. (c) Spread of the ensemble mean. (d) Spread of the ensemble mean of clouds after superposition of centers of mass (cf text) (Fischer et al. (1979)).

proach is to follow each cloud singly, superposing the centers of mass, and then to average the ensemble of releases.

The different averages can be expressed mathematically. Let $n(x,y,t)$ be the concentration observed at point x,y at time t and let $p(n|x,y,t)dn$ be the joint probability that the concentration of tracer material has a value between n and $n + dn$ at the point x,y at time t. Then

$$C(x,y,t) = E\{n(x,y,t)\} = \int_0^\infty np(n|x,y,t)dn \tag{2.8}$$

where E indicates the ensemble mean and $C(x,y,t)dn$ means the ensemble average of the concentration $n(x,y,t)$ after many repeated trials in which identical clouds of particles are released under the same statistical conditions.

For a single cloud the x-coordinate of the center of mass $\underline{x}$ is found as

$$\underline{x} = \frac{1}{M}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xn(x,y,t)dxdy \tag{2.9}$$

where M is the total mass of the cloud. The variance of the cloud, i.e. the mean square x displacement about the center of mass of particles in a single cloud, is given by

$$\sigma_x^2 = \frac{1}{M}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\underline{x})^2 n(x,y,t)dxdy \tag{2.10}$$

The expected x position of the center of mass is

$$E\{\underline{x}\} = \frac{1}{M}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xC(x,y,t)dxdy \tag{2.11}$$

The overall variance of the concentration in the ensemble of clouds with respect to the expected position of the overall center of mass is

$$\Sigma_x^2 = \frac{1}{M}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-E\{\underline{x}\})^2 C(x,y,t)dxdy \tag{2.12}$$

It can be seen that

$$\Sigma_x^2 = E\{\sigma_x^2\} + E\{(\underline{x} - E\{\underline{x}\})^2\} \tag{2.13}$$

or in words: The variance of the ensemble distribution about its expected position is equal to the ensemble average of the variance of each cloud about its center of mass plus the ensemble mean square displacement of an individual cloud's center of mass from its expected position.

The size of an individual cloud can be defined as

$$\ell(t) = (1/2(\sigma_x^2 + \sigma_y^2))^{1/2} \tag{2.14}$$

where in case $\sigma_x = \sigma_y$, $\ell(t) = \sigma_x$ is the radius in a circle. The size of an average cloud is

$$L(t) = (1/2(\Sigma_x^2 + \Sigma_y^2))^{1/2} \tag{2.15}$$

indicating that the average size of a cloud as formed from the ensemble is

$$L^2(t) = E\{\ell^2(t)\} + \frac{1}{2}\left[E\{(\underline{x}-E\{\underline{x}\})^2\} + E\{(\underline{y}-E\{\underline{y}\})^2\}\right] \tag{2.16}$$

confirming the result from Fig. 2.2 where the width of the ensemble mean concentration profile is larger than the average width of a cloud.

Thanks to Taylor's classical work (Taylor, 1921) it is possible to calculate the length L, using some sensible assumptions. Looking at one dimension and only at stationary homogeneous turbulence and using a coordinate system following the mean velocity of the current, it is seen that

$$L^2(t) = \Sigma_x^2 = \frac{1}{M}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 c(x,y,t)\,dx\,dy \tag{2.17}$$

which means that

$$L^2(t) = E\{x^2\} \tag{2.18}$$

Thus, finding the ensemble mean size of the cloud is equivalent to finding the ensemble mean square displacement of the fluid particles.

The location of a particle is

$$x(t) = \int_o^t u dt \quad , \tag{2.19}$$

where u is the velocity and thereby

$$E\{x^2(t)\} = \int_o^t \int_o^t E\{u(\tau_1)u(\tau_2)\} d\tau_1 d\tau_2 \tag{2.20}$$

Now the correlation function is used

$$R_x(\tau_2-\tau_1) = E\{u(\tau_1)u(\tau_2)\}/E\{u^2\} \quad , \tag{2.21}$$

thus giving

$$E\{x^2(t)\} = E\{u^2 \int_o^t \int_o^t R_x(\tau_2-\tau_1) d\tau_2 d\tau_1 \tag{2.22}$$

By changing the variables of integration: $S = \tau_2 - \tau_1$ and $\tau = (\tau_1+\tau_2)/2$, Taylor solved the integral to

$$E\{x^2(t)\} = 2E\{u^2\} \int_o^t (t-s) R_x(s) ds \tag{2.23}$$

For short times ($R_x = 1$) it gives

$$E\{x^2\} = E\{u^2\}\ t^2 \tag{2.24}$$

and for long times

$$E\{x^2\} = 2E\{u^2\}\, T_x t \tag{2.25}$$

where

$$T_x = \int_0^\infty R_x(s)ds \tag{2.26}$$

is the Lagrangian time scale. Also, for long time

$$dE\{x^2\}/dt = 2E\{u^2\}T_x \tag{2.27}$$

showing that after some time the cloud grows linearly with time.

The similarities with molecular diffusion are obvious, and a "turbulent mixing coefficient" can be defined

$$\varepsilon_x = \frac{1}{2}\frac{dE\{x^2\}}{dt} = E\{u^2\}T_x \tag{2.28}$$

relevant some time after the release when the initial separation is forgotten, and leading to a diffusion equation

$$\frac{\partial C}{\partial t} = \varepsilon_x \frac{\partial^2 C}{\partial x^2} + \varepsilon_y \frac{\partial^2 C}{\partial y^2} \tag{2.29}$$

When a Lagrangian length scale ℓ_L is defined by

$$\ell_L^2 = E\{u^2\}T_x^2 \tag{2.30}$$

indicating the distance a particle will travel before losing memory of its initial velocity, it is seen that $\varepsilon_x = \ell_L(E\{u^2\})^{1/2}$, stating that the turbulent mixing coefficient is the product of the Lagrangian length scale and the intensity of the turbulence.

A more general use of ε_x is an expansion of the diffusion equation to

$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} + v\frac{\partial C}{\partial y} = \frac{\partial}{\partial x}(\varepsilon_x\frac{\partial C}{\partial x}) + \frac{\partial}{\partial y}(\varepsilon_y\frac{\partial C}{\partial y}) \tag{2.31}$$

If the time from release is less than the Lagrangian time scale T_x, a careful investigation of relative diffusion, i.e. the separation between two particles, is needed. Now an ensemble mean concentration is formed by averaging the concentration at points equidistant from the center of mass of each cloud in the trial. A new term is now defined, the ensemble mean concentration formed by aligning the centers of mass

$$\psi(\chi,t) = E\{n(X-\underline{X},t)\} \tag{2.32}$$

where

$$\chi = X - \underline{X} \quad , \tag{2.33}$$

and from this

$$E\{\sigma_x^2\} = \frac{1}{M}\int\int \chi_x^2\psi(\chi,t)d\chi_x d\chi_y \tag{2.34}$$

Looking only at two particles separated by the distance s in one dimension and using dimensional analysis, Batchelor (1952) came to the result

$$\frac{dE\{s^2\}}{dt} \sim \varepsilon^{1/3}[E\{s^2\}]^{2/3} \tag{2.35}$$

showing that the rate of increase of the mean square separation of particles is proportional to the mean square separation to the power 2/3. Since the description of the mean square separation from the center of mass of a pair for all pairs of points in an ensemble will be identical and will be exactly the same as the description of the mean square displacement from the center of mass, it can be shown that a differential equation exists

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial x_x}\left(K \frac{\partial \psi}{\partial x_x}\right) \tag{2.36}$$

where

$$K = \frac{1}{2} \frac{dE\{x^2\}}{dt} = \alpha E\{x^2\}^{2/3} \tag{2.37}$$

α being a constant. The diffusion coefficient is proportional to the 4/3 power of the size of the cloud, and this is known as the "4/3-law". In Fischer et. al. (1979) it is said that the differential equation in the ocean can be used from 10^1 m to 10^6 m.

Fig. 2.3 illustrates the influence of the size of the eddies on the dispersion/transport process.

For a more detailed description of the growth of a cloud in a homogeneous and stationary field of turbulence, see e.g. Mikkelsen (1982), which also contains a discussion of the experimental verification of the theories used in the analysis.

2.2. Time series analysis

2.2.1 Introduction

The statistical analysis of the current measurements (Chapter 4 and 5) and the dispersion model described in Chapter 6 are based on time series analyses; therefore an introduction to this topic will be given in the present chapter. The term time series is used to describe a realisation of a stochastic process. Examples of this are legion and exist in many fields: economics, physical sciences, social sciences, geography, etc. A few examples can be stated here: stock prices, temperature measurements, emigration rates, etc.

The intention in this chapter is to provide a basis for the chapters which follows. It starts with a description of Box-Jenkins method for ARMA models, (Box and Jenkins (1976)). Then, in 2.2.3 the use of these processes in model building is de-

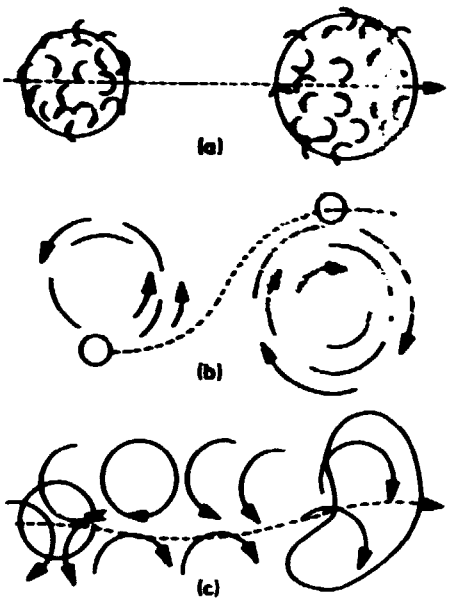


Fig. 2.3. Idealized dispersion patterns. (a) A large cloud in a uniform field of small eddies. (b) A small cloud in a uniform field of large eddies. (c) A cloud in a field of eddies of the same size as the cloud (Slade, 1968).

scribed. In 2.2.4, the spectral representation for a process is investigated and extensions of the theory, such as multivariable processes and adaptive estimation, are presented.

2.2.2. Arima-models

To make the notation easier some operators are used below. The one most often used is the backward shift operator defined by $BZ_t = Z_{t-1}$; thus $B_j z_t = z_{t-j}$. The inverse of this is the forward shift operator $F = B^{-1}$ defined by $FZ_z = Z_{t+1}$; thus $F_j Z_t = Z_{t+j}$. Also important is the backward difference operator $\nabla_Z = Z_t - Z_{t-1} = (1-B)Z_t$ and its inverse, the summation operator,

$$\nabla^{-1} z_t = S z_t =$$

$$\sum_{j=o}^{\infty} z_{t-j} = z_t + z_{t-1} + z_{t-2} + \ldots = (1 + B + B^2 + \ldots) z_t = \tag{2.38}$$

$$(1-B)^{-1} z_t \quad .$$

For discrete time series an important analytical tool is the autocovariance and the autocorrelation function.

The covariance between z_t and its value z_{t+k} separated by k lags (intervals of time) is called the autocovariance at lag k and is defined by

$$\gamma_k = COV[z_t, z_{t+k}] = E[(z_t - \mu)(z_{t+k} - \mu)] \tag{2.39}$$

where μ is the mean value for the process.

The autocorrelation at lag k is defined by

$$\rho_k = \frac{E[(z_t-\mu)(z_{t+k}-\mu)]}{\sigma_z^2} = \frac{\gamma_k}{\gamma_o} \tag{2.40}$$

implying that $\rho_o = 1$.

Used in the identification of an ARMA-model is also the partial autocorrelation coefficient defined by

$$\phi_{kk} = \frac{\det \underline{\underline{U}}_k}{\det \underline{\underline{V}}_k} \quad , \tag{2.41}$$

where

$$\underset{=}{U}_k = \sigma_z^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_{k-3} & \rho_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 & \rho_k \end{bmatrix} \tag{2.42}$$

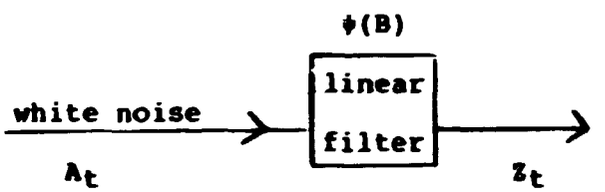
$$V_k = \sigma_z^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_{k-3} & \rho_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 \quad 1 & \end{bmatrix} \tag{2.43}$$

The models described in the following are examples of what are called linear discrete processes. They can be looked upon as output from a linear system to which the input is white noise, i.e. series of independent "shocks" A_t satisfying

$$\mu_t = E(A_t) = 0 \tag{2.44a}$$

$$\sigma_t^2 = V(A_t) = E(A_t^2) = \sigma^2 \tag{2.44b}$$

$$\gamma_k = \mathrm{cov}(A_t, A_{t+k}) = E(A_t \cdot A_{t+k}) = 0, \quad k \neq 0 \tag{2.44c}$$



Thus, the process has the representation:

$$Z_t = \mu + A_t + \psi_1 A_{t-1} + \cdots = \mu + \psi(B) A_t \tag{2.45}$$

where

$$\psi(B) = 1 + \sum_{j=1}^{\infty} \psi_j B^j \ . \tag{2.46}$$

If the process has to be described in terms of earlier values of itself instead of earlier shocks the operator π is needed

$$\pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j \quad , \tag{2.47}$$

thus leading to

$$\tilde{z}_t = \phi(B)A_t \quad \text{and} \quad \pi(B)\tilde{z}_t = A_t \tag{2.48}$$

where

$$\tilde{z}_j = z_t - \mu \tag{2.49}$$

If $\phi(B)$ or $\pi(B)$ are finite series (i.e. ϕ_j or π_j is zero after a certain point) they are called autoregressive (AR) or moving average (MA) processes. A MA(q)-process can be written as follows: (below, z_t is used instead of $\tilde{z}_t$, assuming that the process has zero mean)

$$z_t = A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \ldots - \theta_q A_{t-q}, \tag{2.50a}$$

or

$$z_t = (1-\theta_1 B \ldots -\theta_q B^q)A_t = \theta(B)A_t \tag{2.50b}$$

This process is always stationary but only invertible if the roots in the characteristic equation $z^q - \theta_1 z^{q-1} - \ldots \theta_q = 0$ lie inside the unit circle.

An AR(p)-process can similarily be written as follows:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + A_t \tag{2.51a}$$

or

$$(1-\phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) z_t = A_t \tag{2.51b}$$

or

$$\phi(B) z_t = A_t \tag{2.51c}$$

The process is always invertible but stationary only if all roots in the equation

$$z^p - \phi_1 z^{p-1} - \ldots - \phi_p = 0 \tag{2.52}$$

lie inside the unit circle.

These two types of processes can be combined to form an auto-regressive moving average process of order (p,q).

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \ldots \theta_q A_{t-q} \tag{2.53a}$$

or

$$(1 - \phi_1 B^2 - \ldots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_q B^q) A_t \tag{2.53b}$$

or

$$\theta(B) z_t = \theta(B) A_t \tag{2.53c}$$

where ϕ and θ are polynomes of orders p and q, respectively.

If the original process z_t is not stationary, a "differencing" can be made. If this new process is stationary and can be described by an ARMA-process, the original process is said to be an integrated ARMA or ARIMA-process. An ARIMA (p,d,q) process is defined as follows: $\phi(B)\nabla^d z_t = \theta(B)A_t$, where the difference operator ∇ is described earlier, and where d is the degree of differencing.

In case of seasonal variation in the data, the ARIMA-process can be expanded to a general multiplicative seasonal model. A

seasonal difference operator is defined by $\nabla_s = 1 - B^s$, i.e. $\nabla_s z_t = z_t - z_{t-s}$, and the multiplicative $(p,d,q) \times (P,D,Q)_s$ seasonal model is defined by

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D z_t = \theta(B)\Theta(B^s)A_t \qquad (2.54)$$

where ϕ and θ are polynomiums of degrees p and q and Φ and Θ are polynomiums of degrees P and Q. An example from Conradsen and Spliid (1981) can illustrate the method and the advantage:

A $(2,1,1) \times (1,1,0)_{12}$ model is described by:

$$\begin{aligned}\phi(B)\Phi(B^{12})\nabla z_t = \theta(B)A_t \Rightarrow [1 &- (1+\phi_1)B - (\phi_1-\phi_2)B^2 \\ &+ \phi_2 B^2 - \Phi_1 B^{12} + \Phi_1(1+\phi_1)B^{13} - \Phi_1(\phi_1-\phi_2)B^{14} \\ &- \phi_2\Phi_1 B^{15}]z_t = (1-\theta_1 B)A_t\end{aligned} \qquad (2.55)$$

Using only four parameters, differences up to the order of 15 are constructed. The model was used to describe the tidal variation at Esbjerg on the North Sea Coast. This variation includes an almost sinusoidal component with a 12.42 hours period pretty well approximated by the model which described some 95% of the variation.

2.2.2.1. Numerical examples. Figure 2.4 shows two examples of the autocorrelation function: one for the north component of the current velocity in a specific point and the other for the east component. Both are taken from an earlier investigation of current measuremnts (Boelskifte 1980).

These current velocities were found to satisfy the following ARMA (1,1)-models:

$$V_n(t) = 0.96\ V_n(t-1) - 0.55\ A_n(t-1) + A_n(t) \qquad (2.56a)$$

$$V_e(t) = 0.95\ V_e(t-1) - 0.31\ A_e(t-1) + A_e(t)\ . \qquad (2.56b)$$

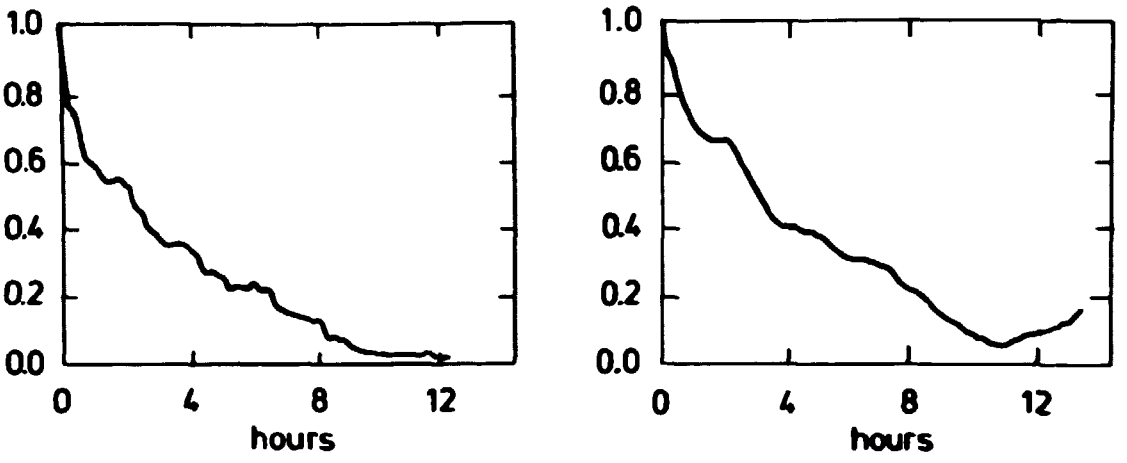


Fig. 2.4. Two examples of an autocorrelation function, north and east component.

The next example is taken from a quite different field of science (Fogerty, 1984); it makes use of a seasonal model and describes the catch, Z_t, of the American lobster

$$Z_t = (1-\theta_1 B)(1-\theta B^{12})A_t$$

$$= (1+0.4317B)(1-0.3417B^{12})A_t \qquad (2.57)$$

$$= 0.4317A_{t-1} - 0.3417A_{t-12} - 0.1475A_{t-13} + A_t$$

2.2.3. Use of ARIMA-models

In the previous section some definitions related to ARIMA-processes have been given. In this section the use of these processes in practice is described. The fitting of an ARIMA-model to some data is a three-step process. First, identification must be performed, i.e. what is p,d,q, or the order of the polynomiums. In Box and Jenkins (1976) identification is described as "the use of the data, and of any information on how the series was generated, to suggest a subclass of parsimonious models worthy to be entertained". Identifying d, the degree of differencing, is the

first step, and here the autocorrelation function can be used, because it will tend to zero for lag $\to \infty$ if the process is stationary. For different values of d the autocorrelation for the process $\nabla^d z_t$ is examined. When it goes towards zero quickly, d is the necessary degree of differencing. Usually d equals 0, 1 or 2 and most often it will be necessary only to inspect the first 20 autocorrelations.

To find the values of p and q, the autocorrelation and the partial autocorrelation are used. For a MA-process the autocorrelation has a cutoff after q lags, while its partial autocorrelation tails off. For an AR(p)-process the partial autocorrelation has a cutoff after p lags and the autocorrelation tails off. If both the autocorrelations and partial autocorrelations tail off, a mixed process has to be identified. If $q > p$ the autocorrelation function is a mixture of exponentials and damped sine waves after the first $q - p$ lags. If $p > q$ the partial autocorrelation function is dominated by a mixture of exponentials and damped sine waves after the first $p - q$ lags. Examples of identification and handling of problems due to large variances in and high autocorrelation between estimated autocorrelations can be found in Box and Jenkins (1976).

Estimates of the parameters can be made in many ways. However, it is not the intention here to discuss different procedures. Again a reference is made to Box and Jenkins (1976) where the estimate is examined in detail. In the present study a method developed by Spliid (1983) is used. It is a fast estimate algorithm designed to handle multivariate ARIMA-models (see Ch. 5).

After the estimate, diagnostic checks are applied to the fitted model. The principle is to look at the residuals, i.e. the deviation of the original process from the estimated one, and see if they fulfil some restrictions. Since the estimation is a quick procedure on a computer, in practice one fits different models (different p's and q's) to the process, and then compares the properties of these models and their residuals. Examples of this will be given in Chapter 5.

One of the advantages of ARIMA models is that forecasting is a relatively simple process. The calculations are based on minimum mean square error forecasts, i.e. the principle is to minimize the term $E[(Z_{t+\ell} - \hat{Z}_t(\ell))^2]$, where $\hat{Z}_t(\ell)$ is the forecast at time $t+\ell$. For an ARIMA-process $\phi(B)\nabla^d Z_t = \theta(B)A_t$ the term

$$\hat{Z}_t(\ell) = \psi_\ell A_t + \psi_{\ell+1}A_{z-1} + \psi_{\ell+2}A_{t-2} + \cdots \tag{2.58}$$

is the forecast after minimizing the expected mean square error. The $\psi(B)$ polynomium can be found by solving the equation system system

$$\begin{aligned}
&\psi_j = 0 \qquad j < 0 \\
&\psi_o = 0 \\
&\psi_1 = \phi_1 - \theta_1 \\
&\psi_2 = \phi_1\psi_1 + \phi_2 - \theta_2 \\
&- \\
&- \\
&\psi_j = \begin{cases} \phi_1\psi_{j-1} + \cdots \phi_{j-1}\psi_1 + \phi_j - \theta_j & , \quad j \leq p + d \\ \phi_1\psi_{j-1} + \cdots \phi_{p+d}\psi_{j-p-d} - \theta_j & , \quad j > p + d \end{cases} \\
&\theta_j = 0 \quad \text{for} \quad j > q
\end{aligned} \tag{2.59}$$

Another important property of ARIMA-models is the relative ease with which they allow simulations of a system to be created. Examples of this are given in Chapter 6.

Some expansion of the ARIMA-processes are the transfer function models used to describe dynamic systems with a known input X_t, an output Y_t and if necessary some added noice N_t, where $Y_t = v(B)X_t$, and where $v(B)$ is called the transfer function of the system. If the noise is taken into account, it is assumed that it can be represented by an ARIMA (p,d,q) process

$$N_t = \phi^{-1}(B)\theta(B)A_t, \tag{2.60}$$

leading to the model

$$Y_t = \delta^{-1}(B)\omega(B)X_{t-b} + \phi^{-1}(B)\theta(B)A_t \tag{2.61}$$

given

$$\delta(B)Y_t = \omega(B)X_{t-b} \tag{2.62}$$

and

$$v(B) = \delta^{-1}(B)\omega(B)B^b \tag{2.63}$$

These types of models are more general than conventional regression analysis. They describe a relationship between two physical parameters in a dynamic system disturbed by noise. They are also able to compensate for the effects of trend, seasonal variation and other nonstationary behaviour. Hosmer (1984) provides a significant oceanographic example in which water temperature measurements are analysed in order to describe the current situation in a large area.

Water temperature recordings from the northeastern US continental shelf at different positions and different depths during the years 1950-1980 have been analysed. The region has been divided into 11 compartments, and relations between temperature measurements in these have been sought out. Seasonal variation could be modelled. A transfer function model describes the monthly (13 lunar months per year) temperature in the area called Georges Bank shallow area (X) as a function of the temperature in the Middle Atlantic deep area (water depths greater than 150 meters) (Y):

$$\begin{aligned} X_t = {} & X_{t-1} + X_{t-13} - X_{t-14} + 0.24Y_{t-2} - 0.24Y_{t-3} - 0.24Y_{t-15} \\ & + Y_{t-16} + a_t - 0.82a_{t-1} - 0.75a_{t-13} + 0.62a_{t-14} \end{aligned} \tag{2.64}$$

Similar models are estimated for the other areas, and a description of the current between the compartments is found to agree with oceanographic data. The paper concludes that "bivariate ARIMA models could be used as another tool for unraveling some of the mysteries of ocean dynamics".

It is expected that in the future, time series analysis will play a more dominant role in the statistical analysis of environmental data. To quote Austin et. al. (1984) "Time series analysis offers the basis for a bridge between the observations and model building".

The bivariate models lead to an important feature widely used in the present study, viz. the cross correlation function. The series Z_t and Y_t can be looked upon as one bivariate stochastic process, and the cross covariance $\gamma_{xy}(k)$ between the two series separated by k lags is given by

$$\gamma_{xy}(k) = E[(x_t-\mu_x)(Y_{t+k}-\mu_x)] \quad , \quad k = 0,1,2 \ldots \qquad (2.65)$$

and similarly for γ_{yx}.

It should be noticed that $\gamma_{xy}(k)$ generally differs from $\gamma_{yx}(k)$, but that $\gamma_{xy}(k) = \gamma_{yx}(-k)$. The process is not necessarily symmetrical about k = 0. From this, the cross correlation function is defined:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x\sigma_y} \quad , \quad k = 0, \pm 1, \pm 2 \ldots \qquad (2.66)$$

2.2.4. Spectral representation of time series

Until now, the series have been studied in the time domain. Another approach is to look at their representation in the frequency domain. The sample spectrum of a time series is the cosine transform of the estimate of the autovariance function (Box & Jenkins, 1976).

$$I(f) = 2(c_o + 2 \sum_{k=1}^{N-1} C_k \cos(2\pi fk)) \quad , \quad 0 < f < 1/2 \qquad (2.67)$$

where N is the length of the series and C_k the estimate of the autovariance function. The power spectrum is defined by

$$\Gamma_{xx}(f) = \Delta \sum_{k=-\infty}^{\infty} \gamma_{xx}(k) e^{-i2\pi kf\Delta} \quad , \quad \frac{1}{2\Delta} < f < \frac{1}{2\Delta} \qquad (2.68)$$

Δ is the time interval.

A detailed description of spectra and their use is found in (Jenkins & Watts, 1968). When using filters (highpass, lowpass, and bandpass), the spectral repesentation is the most convenient. In interpreting some physical measurements, spectra are also very useful, as will be seen in Chapter 4. (Also, Koopmans (1974) has a comprehensive description of spectral analysis).

When a signal has the main part of its variance at a few frequencies (e.g. water level measurements have a peak around the tidal cycle, approx. 12.42 h) this feature disturbs the analysis at other frequencies. To filter out these frequencies a variety of filters are available. The principle applied here is a "window" following the time series and performing one or another kind of integration:

$$Y(t) = \int_{-T}^{T} h_\ell(u) X(t-u) du \qquad (2.69)$$

where $h_\ell(t)$ is the transfer function. The impulse response function (or weight function) of the filter is the transfer function multiplied by the actual "window".

For an ARMA model $\phi(B)Zt = \theta(B)A_t$ the spectrum is

$$S(f) = \frac{\sigma^2}{2\pi} \left| \frac{\sum_{k=0}^{p} \phi_k e^{-ifk}}{\sum_{j=0}^{q} \theta_j^{-ifj}} \right|^2 \qquad (2.70)$$

Thus the ARMA-model can be seen as a parameterized description of the process, whereas the spectrum is a function of the frequency f.

2.2.5. Other ARIMA models

The bivariate stochastic process has already been mentioned and a natural expansion of that is the multivariate (or vector) stochastic process. For a multivariate ARMA(p,q)-process of dimension r the representation is as follows:

$$\underline{Z}_t = \underline{\underline{\phi}}_1 \underline{Z}_{t-1} + \ldots + \underline{\underline{\phi}}_p Z_{t-p} + \underline{A}_t - \theta_1 \underline{A}_{t-1} - \ldots - \underline{\underline{\theta}}_1 \underline{A}_{t-q} \qquad (2.71a)$$

or in general

$$\underline{\underline{\phi}}(B)\underline{Z}_t = \underline{\underline{\theta}}(B)\underline{A}_t \qquad (2.71b)$$

where $\underline{Z}_t$ and $\underline{A}_t$ are vectors of dimension r and $\underline{\underline{\phi}}_k$ and $\underline{\underline{\theta}}_k$ are matrices of dimension r x r. As an example, the resulting equations are given below for r = 2, p = 1 and q = 1

$$\underline{Z}_t = \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix}, \underline{A}_t = \begin{pmatrix} A_{1,t} \\ A_{z,t} \end{pmatrix} \qquad (2.72a)$$

$$\underline{\underline{\phi}} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \text{ and } \underline{\underline{\theta}} = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \qquad (2.72b)$$

leading to

$$z_{1,t} = \phi_{11}z_{1,t-1} + \phi_{12}z_{2,t-1} + A_{1,t} - \theta_{11}A_{1,t-1} - \theta_{12}A_{2,t-1} \tag{2.72c}$$

$$z_{2,t} = \phi_{21}z_{1,t-1} + \phi_{22}z_{2,t-1} + A_{2,t} - \theta_{21}A_{1,t-1} - \theta_{22}A_{2,t-1} \tag{2.72d}$$

$\underline{A} \in N(\underline{0}, \underline{\underline{\Sigma}})$ is white noise.

For multidimensional time series an important concept is the covariance matrix defined by

$$\underline{\underline{\Sigma}} = \begin{vmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ - & & & \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{vmatrix} \tag{2.73}$$

where σ_{12} is the covariance between z_1 and z_2.

Estimates of MARIMA-processes can be made using the algorithm by Spliid (1983) mentioned previously. After making the estimation the presence of cross correlation in the residuals must be tested. This can easily be done as follows:

$$\text{Model: } \hat{\underline{\underline{\phi}}}(B)\underline{Y}_t = \hat{\underline{\underline{\theta}}}(B)\underline{\hat{A}}_t \tag{2.74}$$

To test whether the model fits the data, yet another model can be fitted to the residuals,

$$\underline{\underline{\xi}}(B)\underline{A}_t = \underline{E}_t \tag{2.75}$$

where $\underline{\underline{\xi}}(B)$ is an autoregressive matrix polynomium of relatively high order. If $\underline{\underline{\xi}}$ turns out to be significantly different from the unity matrix, this indicates that the ARMA-model does not describe all important aspects of the original data. The significance of each parameter in $\underline{\underline{\xi}}$ can be checked using an F test technique ($F = (\hat{\phi}_p/\hat{\sigma}_{\phi p})^{2)}$. See also Spliid(1980).

2.2.6. The Kalman filter

In many applications adaptive forecasting is an important tool. Therefore, it will be described briefly as follows: Consider the discrete system described by the equations

$$X(t+1) = \underline{\underline{\phi}}X(t) + v(t) \tag{2.76}$$

$$Y(t) = \underline{\underline{\theta}}X(t) + e(t) \tag{2.77}$$

where X is an n-dimensional state vector, Y a p-dimensional vector of observed outputs and v(t) and e(t) are sequences of independent Gaussian vectors with zero mean values and the covariances

$$\begin{aligned} Ev(t_1)v^T(t_2) &= R_1 \quad \text{for } t_1 = t_2 \text{ , } 0 \text{ otherwise} \\ Ev(t_1)e^T(t_2) &= 0 \\ Ee(t_1)e^T(t_2) &= R_2 \quad \text{for } t_1 = t_2 \text{ , } 0 \text{ otherwise} \end{aligned} \tag{2.78}$$

and where the matrices $\underline{\underline{\phi}}$, $\underline{\underline{\theta}}$, $\underline{\underline{R}}$, and $\underline{\underline{R}}_2$ may depend on time. The description is based on the one given by Åstrøm (1970), but a number of textbooks give an introduction to the Kalman filter, e.g. Bennett (1979). The estimate of the state at time t + 1 based on earlier outputs from the system which minimizes the prediction error is the conditional mean $\hat{X}(t+1|t)$ which satisfies the recursive equation

$$\hat{X}(t+1|t) = \phi\hat{X}(t|t-1) + K(t)[Y(t) - \theta\hat{X}(t|t-1)] \tag{2.79a}$$

$$\hat{X}(t_o|t_o-1) = m \tag{2.79b}$$

The matrix K(t) is given by

$$K(t) = \phi P(t)\theta^T[\theta P(t)\theta^T+R_2]^{-1} \tag{2.80}$$

where P(t) is the covariance of the estimation error

$$P(t+1) = \phi P(t)\theta^T + R_1 - \phi P(t)\theta^T[\theta P(t)\theta^T+R_2]^{-1}\theta P(t)\phi^T \tag{2.81a}$$

$$P(t_o) = R_o \tag{2.81b}$$

This is called the Kalman filter and has been used for a wide range of problems. Further details on this topic will not be given here. However, in Chapter 5 we will demonstrate its use.

2.3. Some elements of health physics

2.3.1. Introduction

After the introduction to statistical turbulence theory and time series analysis, elements of health physics will be described in this section. Since the purpose of creating dispersion models for radioactive material is to provide the basis for reliable dose assessments, the terms related to the calculation of detriments will be quoted. The dose terms reproduced will be used in the chapter about dispersion models (Ch. 3) and the chapter about the verification of the simulation model (Ch. 7). One reference, (ICRP, 1977), form the basis for the present section.

2.3.2. Some definitions

Like energy, activity is an example of a word that has a colloquial meaning and a different, precise scientific meaning. The radioactivity, A, is defined by

$$A = dN/dt \tag{2.82}$$

where dN is the expectation value of the number of spontaneous nuclear transitions in the time interval. The unit is Becquerel [Bq], giving the number of transitions per second. (Old unit: 1 Curie = $3.7 \cdot 10^{10}$Bq).

Each nuclear transition results in some kind of radiation, the effect of which on matter depends on the magnitude of the radiation field and on the degree of interaction between the radiation and matter. This is measured by the specific energy

$$z = \varepsilon/m \tag{2.83}$$

where ε is the energy imparted and m is the mass. The unit has a special name, Gray (Gy). 1 Gy = 1 J kg^{-1} (1 Gy = 100 rad). The absorbed dose, D, is derived from the specific energy:

$$D = d\varepsilon/dm \tag{2.84}$$

where $d\varepsilon$ is the mean energy imparted to matter of mass dm by ionizing radiation. The unit is 1 Gy = 1 J kg^{-1}.

Years ago, the unit used in radiation protection was the Röntgen (R). It is the unit for exposure, X, which is the quotient of dQ by dm, where the value of dQ is the absolute value of the total charge of the ions of one sign produced in air when all the electrons liberated by photons in air of mass dm are completely stopped in air

$$X = dQ/dm \tag{2.85}$$

The unit is 1 C/kg (1R = $2.58 \cdot 10^{-4}$C kg). In practical health physics 1R is often considered as the level of exposure which in the body leads to an absorption of energy equivalent to 10 m Gy (1 rad).

The biological effect, and thereby the detriment, of a given absorbed dose depends on many factors. This is expressed by the dose equivalent, H, which is defined as the product of D, Q and N at the point of interest

$$H = DQN, \tag{2.86}$$

where D is the absorbed dose, Q is a quality factor which depends on the kind of radiation, and N is the product of all other modifying factors (usually N equals one). The special unit for H is Sievert ($1\ Sv = 1\ Jkg^{-1}$). Earlier $1\ rem = 10^{-2}\ Sv$ was used (röentgen equivalent man).

ICRP recommends that the principle be used that risk should be considered equal, whether the whole body is irradiated uniformly, or whether there is non-uniform irradiation. This feature is introduced by the definition of the effective dose equivalent

$$H_E = \Sigma_T w_T H_T \tag{2.87}$$

where w_T is a weighting factor representing the proportion of the stochastic risk resulting from tissue (T) to the total risk, when the whole body is irradiated uniformly, and H_T is the annual dose equivalent in tissue (T). The weighting factors are given in (ICRP, 1977).

The relationship between the detriment and distribution of dose equivalent in an exposed population is not simple, but valuable use can be made of the collective dose equivalent

$$S = \Sigma_i H_i P_i \tag{2.88}$$

where H_i is the per caput dose equivalent in the whole body or any specified organ or tissue of the P_i members of subgroup (i) of the exposed population.

From this definition the collective effective dose equivalent, S_e, can also be found, simply by incorporating the weight factors. The collective effective dose equivalent rate, $\dot{S}_E$, can be defined as:

$$\dot{S}_E = dS_E/dt \tag{2.89}$$

Further quantities are needed when the exposure is extended in time. The dose equivalent commitment, H_C, from a given practice, is the infinite time integral of the per caput dose equivalent rate, $\dot{\bar{H}}(t)$, in a given organ or tissue for a specified population:

$$H_C = \int_0^\infty \dot{\bar{H}}(t)dt \qquad (2.90)$$

The collective effective dose equivalent commitment S_E^C is found by integrating the collective effective dose equivalent rate S_E over all time:

$$S_E^C = \int_0^\infty \dot{S}_E(t)dt \qquad (2.91)$$

This gives a measure of the radiation-introduced health detriment to a population over all time.

A special case of the dose equivalent commitment is the committed dose equivalent, H_{50}, to a given organ or tissue from a single intake of radiactive material. This is the dose equivalent that will be accumulated over 50 years:

$$H_{50} = \int_{to}^{t_o+50y} \dot{H}(t)dt \qquad (2.92)$$

where $\dot{H}(t)$ is the relevant dose-equivalent rate and t_o is time of intake.

2.3.3. Radioecology

That field of health physics which deals with radioactivity in the environment is called radioecology. Aarkrog (1979) writes that "this science includes the movement of radionuclides within ecological systems and their accumulation within specific ecosystem components such as air, water, soil and living organisms." The reference contains a variety of radioecological studies and thus gives a detailed view of the whole field. For further introduction reference can be made to e.g. Aarkrog et al. (1982).

3. MODELS AND SCALES

3.1. Introduction

Before a model is constructed, the most important issues to address are the time and length scales that the model must cover. One must determine whether it is the behaviour of the discharge during the initial mixing that is important, or the long-term dilution over, e.g. thousands of kilometers? In Table 3.1, taken from Fischer et. al. (1979), examples of typical scales are shown along with a description of the associated physical process. Since the processes involved in each phase are completely different, they require separate models, even though there is an overlap of the dispersion and transport mechanisms from one scale to another. These different processes may lead either to complicated models covering more than one phase of transport or to restrictive assumptions for each model.

Along with the question of scales, the purpose of the model must be stated. Concerning a fixed scale, several questions can be raised: What will the steady-state situation look like? (i.e. what is the average transport and dilution for e.g. a whole year where the discharge is approximately constant?) What is the possibility that an accidental release, within a given time and within some concentration limits, will reach a specific point? Are the collective doses or individual doses most important (for an explanation of the different doses see Section 2.3) or, in other words, are the mean values over a large area or peak values within a smaller area the most important?

Also, the form of the source will influence the model: Is it a point, a line, or perhaps, as is the case for fallout, an area? (Fallout is the radioactive isotopes deposited after atmospheric nuclear test explotions). Is it a momentaneous or a continous source? All these questions have to be answered before a model is constructed. Various models incorporate different physical parameters, such as the wind speed, topografy, sedimentation rates

etc. Some models describe evolution in time, others a steady-state situation.

Phase	Phenomenon	Length scale[a] (m)	Time scale[a] (sec)
(1)	Initial jet mixing (rise of buoyant jets over an outfall diffuser in a stratified fluid)	$<10^2$	$<10^3$
(2)	Establishment of sewage field or cloud, travelling with the mean current; lateral gravitational spreading.	10^1-10^3	10^2-10^3
(3)	Natural lateral diffusion and/or dispersion.	10^2-10^4	10^3-10^5
(4)	Advection by currents (including scales of water motion too large compared to sewage plume to be called turbulence).	10^3-10^5	10^3-10^6
(5)	Large-scale flushing (advection integrated over many tidal cycles); up- or downwelling; sedimentation.	10^4-10^6	10^6-10^8

[a]) Approximate orders of magnitude

Fig. 3.1. The Effluent Flow from a Sewer Passes through a Succession of Physical Processes at Scales from Small-to-Large (From Fischer et al. (1979)).

Concerning radioactivity from Barsebäck, different models are needed:

a) a model for the steady-state situation in the Danish straits for routine discharges (for monitoring purposes)

b) a model calculating the dose to a critical group, also for routine discharges
c) a model calculating the possible collective dose following an accident
d) a model for the initial mixing in general (e.g. temperature rise)
e) a model able to track a leakage (also relevant if, e.g. oil was considered).

The following sections are brief descriptions of some of the models found in the literature. They cover a span of length and time scales and deal with some of the purposes a)-e). The review, of course, is not at all a complete collection of models nor a comprehensive description of each model. Our intention is to present some ideas and ways of handling the problems.

3.2. Large-scale models

3.2.1. Compartment models

A model for transport in the "northern European waters" has been developed by NRPB (Clark et al., 1980). The area is divided into 17 boxes (Fig. 3.1). Following a discharge into box i, the concentration of a given radionuclide in that box can be determined using the equation

$$\frac{dN_i}{dt} = \sum_{j=1}^{N} (K_{ji}N_j - K_{ij}N_i) + Q_i - \lambda_i^S N_i - \lambda N_i \ , \qquad (3.1)$$

where N_i is the amount of activity in box i
K_{ij} the water exchange rate from box j to box i
Q_i the discharge to box i
λ the decay constant
λ_i^S the sedimentation rate for a given isotope in box i

This model assumes total and instantaneous mixing in each box for every timestep. Data for the water exchange, the volume of the

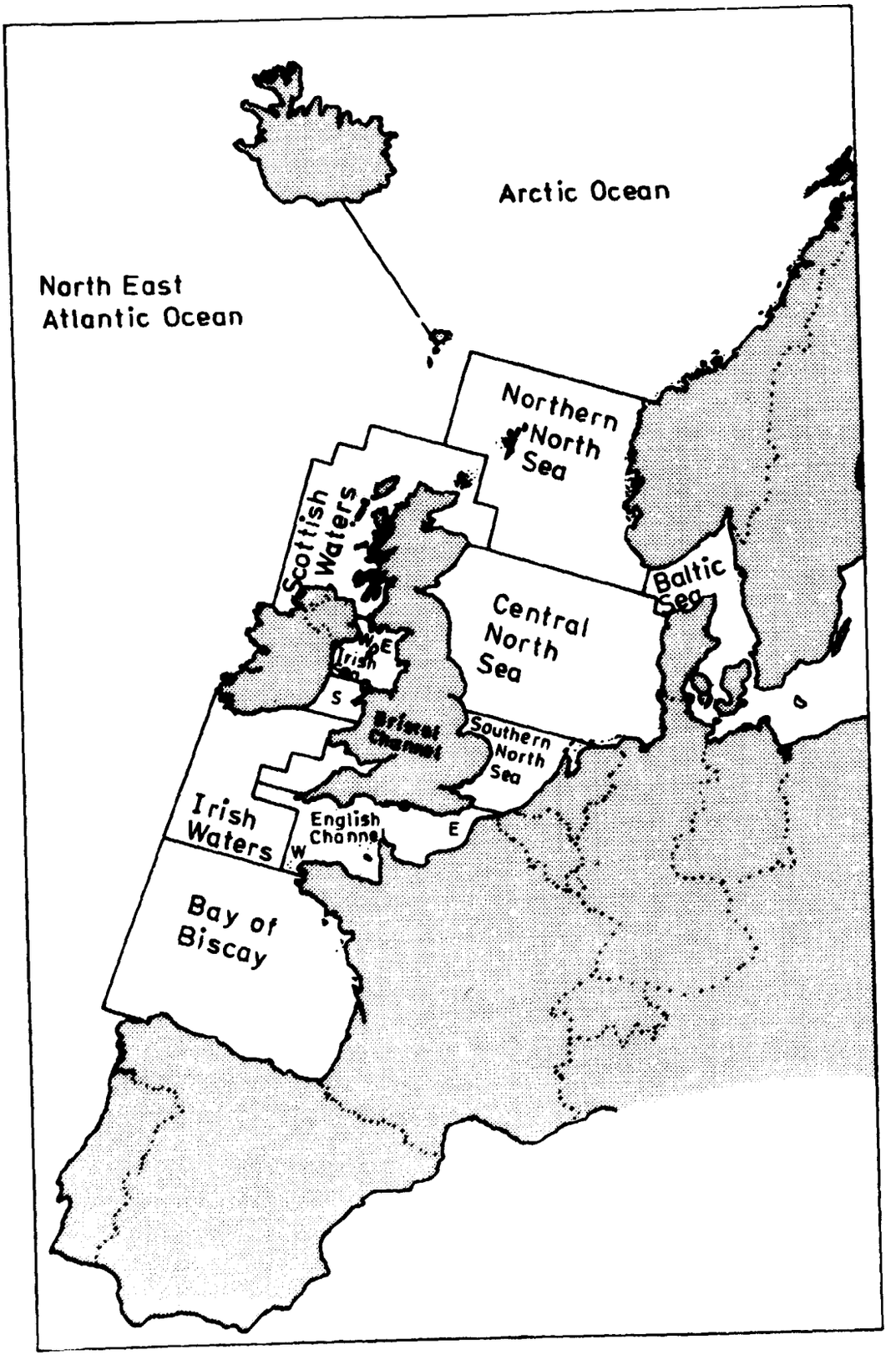


Fig. 3.1. Regional marine compartments for northern European waters.

boxes and the sedimentation rates (dependent on box and isotope) are reported. The purpose of the model is to estimate the collective effective dose equivalent commitment to the population resulting from a unit release of a given radioisotope into any of the boxes. In the dose calculation part, data for seafood catch and concentration factors between fish and water are used. Tests show the model's applicability in describing transport over large length scales. The transport of Cs from Sellafield (Irish Sea) is well described by the model. For example, the 4-years transport time from Sellafield to the Danish Straits (Aarkrog et al., 1982) is confirmed. But the model is not detailed for the Danish straits and therefore useless for that area. While this model can be useful for dose calculation, some of the assumptions are too rigid for the model to provide a detailed description of the transport. Some of the problems are: The water exchange between two boxes depends on time. Total mixing within a box does not take place. The boxes are too big. But to a certain degree these are minor drawbacks compared with the uncertainty of the data used.

Evans (1985) improved the model to make it usable for the Baltic Sea as well. The major change was the subdivision of some boxes into two layers. This is important for the Baltic, where stratification is a pronounced phenomenon. A new approach to compartment models using "backward estimation" of transfer coefficients is used by Hallstadius et al. (1986).

The improved model cannot describe the situation in the Sound in detail, where the exchange rate is high and no data for making a subdivision is accessible. To describe the situation here, other methods must be used. The Sound is a part of the transition area between the Baltic Sea with 25 y mean residence time and the North Sea with approximately 2 y mrt. (these are the figures used e.g. by Clark et al., (1980)) If one uses the box model principle, one finds large uncertainties in calculating the dose from an accidental release from Barsebäck, due to the dependence of the dilution on the current direction at the moment of release, as discussed in Boelskifte (1983).

3.2.2. Empirical models

A general description of the transport of radionuclides released from Sellafield has been possible for Cs and Sr (Aarkrog et. al. 1983). Thanks to very low detection limits for γ-emitting radionuclides, ^{137}Cs can be followed from the Irish Sea, into the North Sea, a part in the Danish straits and the Baltic, along the Norwegian coast, and along the east coast of Greenland. This route is confirmed by measurements of ^{99}Tc in seaweed (Aarkrog, 1982), which illustrate the possibility of using radionuclides in double tracer experiments. Comparing the levels found in water with discharge data and, as far as possible, using the shorter-lived ^{134}Cs, estimates can be given for the dilution and transport times from the source to every other place along the route (1-10y) (see, for instance, Dahlgaard et al. (1984)). (Fig. 3.2)

The value of such an empirical model is that it can describe the transport and dilution of any water soluble pollutant released somewhere along the route, e.g. in the North Sea. Most remarkable is the large scale that it covers (10000 km).

Another empirical description of the transport of radionuclides is given by Mattsson et. al. (1980). Here the measurement of different isotopes in seaweed along the Swedish westcoast leads to a description of the concentration C of a radioisotope in seaweed related to the distance, X, from Barsebäck:

$$C(X) = \alpha X^{-\beta} \qquad (3.2)$$

where α depends on the release and isotope and β is estimated to 1.4. According to the authors the expansion of a "cloud" near the outlet will be proportional to X^3. A vertical limit is rapidly reached at the bottom or the halocline (10-15 m). Thereafter the cloud expands at a rate proportional to the distance squared (β= 2). If the speed of expansion is reduced, β will be smaller, leading to β =1.4.

Some of the problems in this analysis have been studied further in an investigation described in Chapter 7. The problems were,

among others: too few samples were taken, the realibility of seaweed as an indicator for the radionuclide concentration in water was not discussed, and no samples were taken south of Barsebäck

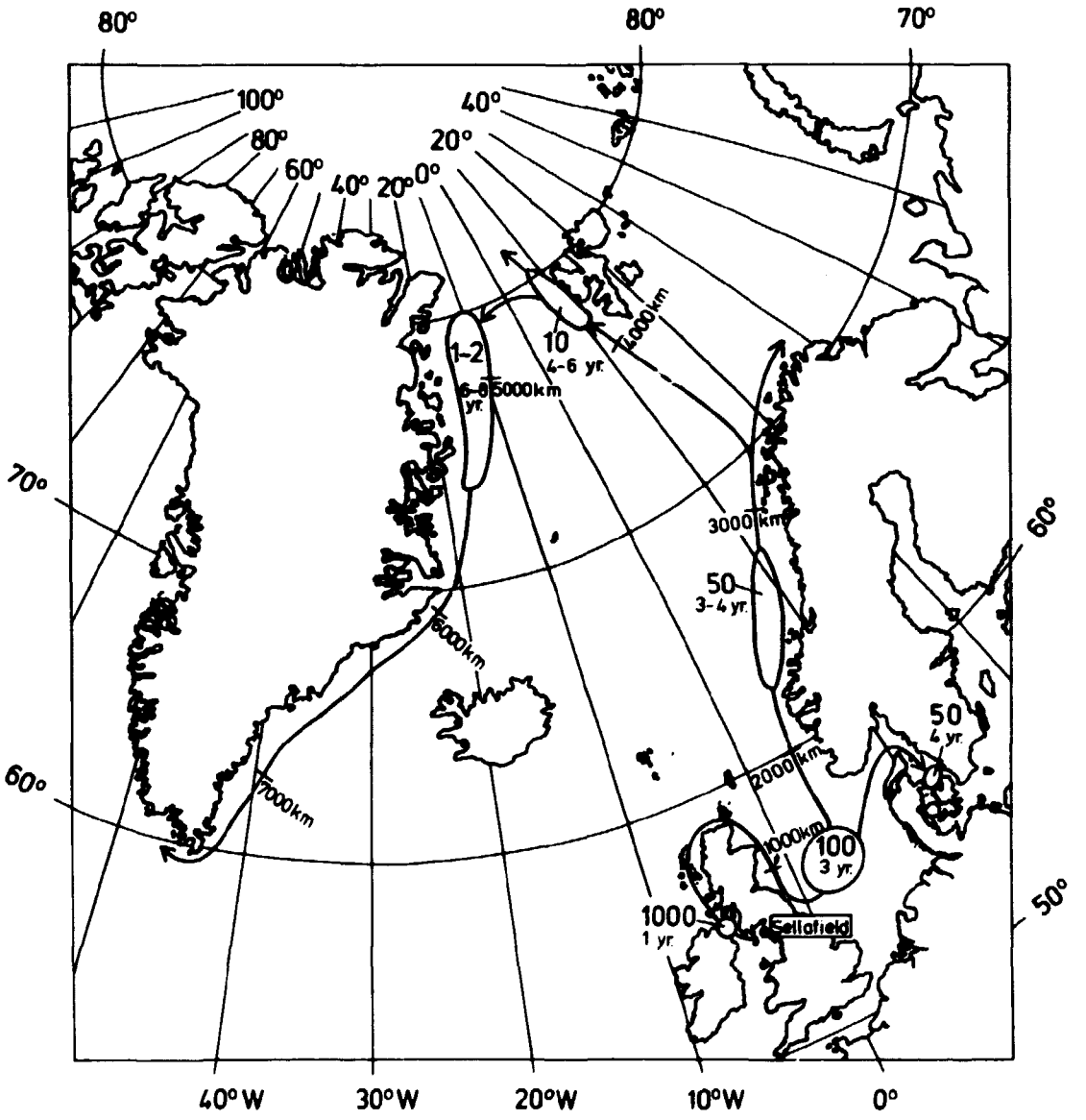


Fig. 3.2. Approximate route for transport of radionuclides from Sellafield to East Greenland. Relative concentrations (~ Bq ^{137}Cs m^{-3}) and approximate transit times are indicated. Distance along the track are shown in km. The relative concentrations can also be read as percent of the concentration in the North Sea of any pollutant behaving approximately conservatively. (Dahlgaard et al. (1984)).

and in the Danish part of the Sound. Furthermore, no lower limit of validity of the equation is given ($X^{-1.4}$ tends towards infinity as X tends towards zero).

As described in Chapter 7 or by Boelskifte (1985), some of the questions have been answered, and some still remain to be investigated.

3.3. "Diffusion constant" models

The theoretical background of turbulent dispersion has been stated in Chapter 2. A significant inspiration in the development of dispersion theories has been the work by Hinze (1959), where a comprehensive description of problems related to turbulent flow is presented. The differential equation for the diffusion from a source in a uniform flow is:

$$\underline{U}_i \frac{\partial P}{\partial X_i} = \varepsilon \frac{\partial^2 P}{\partial X_i \partial X_i} \qquad (3.3)$$

where P is a probability, equivalent to a concentration; $\underline{U}_i$ is the mean velocity (in X_i direction), and ε is the diffusion constant. Some solutions to Eq. (3.3) are given, for instance for a point source:

$$P(X_1,X_2,X_3) = \frac{S}{4\pi r\varepsilon} \exp(-\underline{U}_1(r-X_1)/2\varepsilon)$$

where S is the source term and $r^2 = X_i X_i$.

3.3.1. Sandia model

Practical examples of the theory applied to radionuclides can be found in Niemczyk et. al. (1981). The general equation is the following

$$\frac{dC}{dt} = - U_x\frac{dc}{dx} - U_y\frac{dc}{dy} - U_z\frac{dc}{dz} + \frac{d}{dx}\left[D_x\frac{dc}{dx}\right] + \frac{d}{dy}\left[D_y\frac{dc}{dy}\right] + D_z\frac{dc}{dz}\Big] - \lambda c + S(t) \qquad (3.4)$$

where λ is the decay constant for the radionuclide.

With convenient assumptions the equation is solved for different flow systems. Thus, deterministic models for estuaries, rivers, and the Great Lakes (in USA) have been developed. These areas can be subdivided again into near-shore, surface mixed-tank, complete mixed-tank, etc. The assumptions deal with straight coastline, constant coefficients, etc. The analytical solutions to Eq. (4) consist of exponential functions and depend on time, space and diffusion coefficients. A quantitative solution needs a lot of data, and for the Great Lakes these include volumes, outflow rates, residence times, areas, maximum and mean depths, etc. The diffusion constant D_x ranges from 100 to 500 cm^2/s, and D_y from 1 to 3 cm^2/s. The purpose of this model is to provide data for dose calculations for a steady-state situation, and for a continous release. This kind of model represents many similar models based on the same method, and used in a variety of different areas.

3.3.2. Danish Hydraulic Institute

Another model based on the same equation (Eq. (4)) is described in DHI (1977). It uses a grid structure for the area under consideration and the equations are solved numerically at each point.

The system has been used for simulations on time scales less than a week, and length scales 0.1-100 km or even longer. A lot of input data is needed, current data, dispersion coefficients, etc. Since the calculation method is very complex, the model needs a lot of computer time and space. It is designed to handle dynamic situations, and has been used on a variety of problems in many parts of the world. However, the model is not constructed for the purpose of dose calculations.

3.3.3. IC model

Another application of Hinze's formula was performed by the Danish Isotope Centre (IC) (1973). IC has done many recipient investigations, and for this one, tracer experiments were carried out (see Fig.3.3) to estimate the diffusion coefficient in the actual area. A radioactive tracer (^{82}Br) was released at a single point and was followed for 6 hours, using a boat with radioactivity detectors and data collection equipment. The mean currents for five-minute periods were used as the basis for a current statistic. For different current situations a dilution course is determined following Hinze's formula

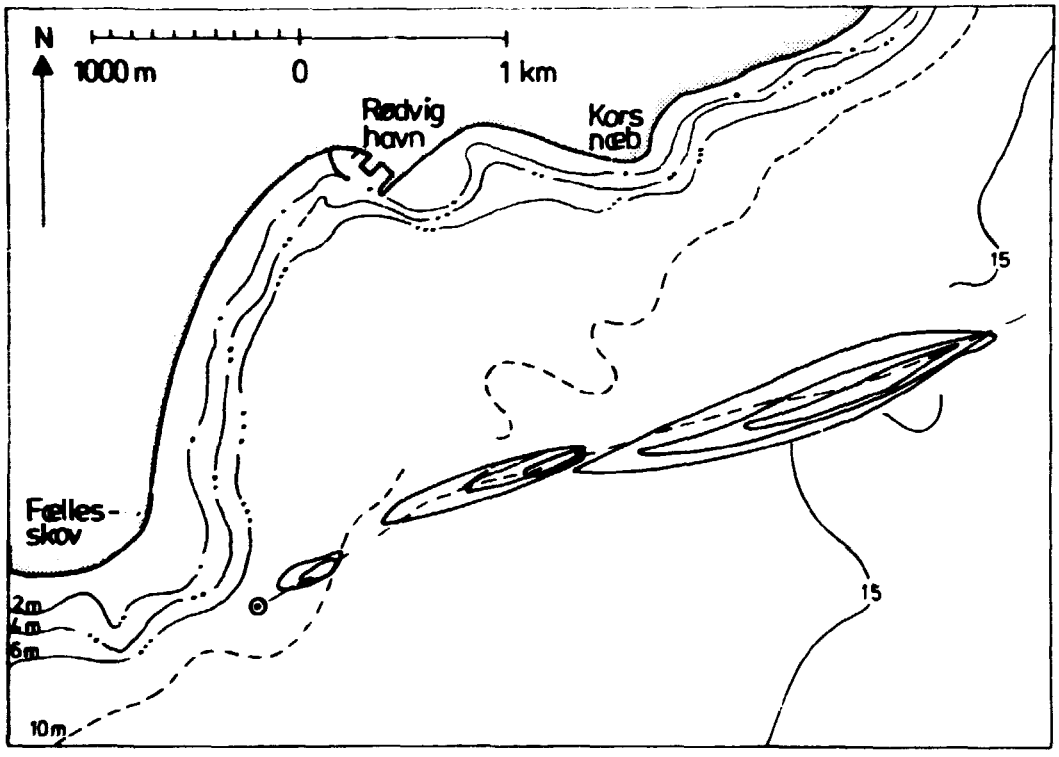


Fig. 3.3. A tracer experiment (IC. 1973).

$$P = \frac{C}{C_o} = \frac{S}{h\sqrt{4\pi DuX}} \exp(uy^2/(4XD)) \qquad (3.5)$$

where h is the depth of the release of the tracer, D the horizontal diffusion coefficient, S the source term, u the velocity, and

C the concentration. The time and length scale covered by this calculation was 6 hours and 5 km. The purpose of the model was to predict the effects on water quality if a sewage outlet were built in the area. Some questions needed to be answered: What was the mean strain on the water quality, with what probability would a high concentration reach the coast or, in other words, how far from the coast would the outlet have to be built to prevent the coast from being polluted, and what would be the dilution in a critical situation with on-shore wind.

Many similar investigations with almost the same purpose have been performed in Denmark and elsewhere. It was the intent of this short description to give an idea of the methods and the problems being solved. Such methods are useful in solving actual problems; however, they are also time consuming. They combine deterministic models with statistical observations. The statistical observations are the background for deciding what current situation should be studied in detail.

3.4. Stochastic simulation models

3.4.1. SMHI model

A combination of a deterministic and a stochastic model is found in Bork (1977). It is applied on the lake Vänern in Sweden (1977) where 10-minute current measurements are made. In the simulation scheme single particles are studied each having a velocity expressed as

$$\frac{dx}{dt} = U + aP_u \qquad (3.6)$$

(equal for Y and Z direction) where U is the mean velocity, a a random number between -1 and +1, and P_u a known turbulence scale. P_u is found from the diffusion constant, A, in this way

$$P_u = (6A/\Delta T)^{1/2} \qquad (3.7)$$

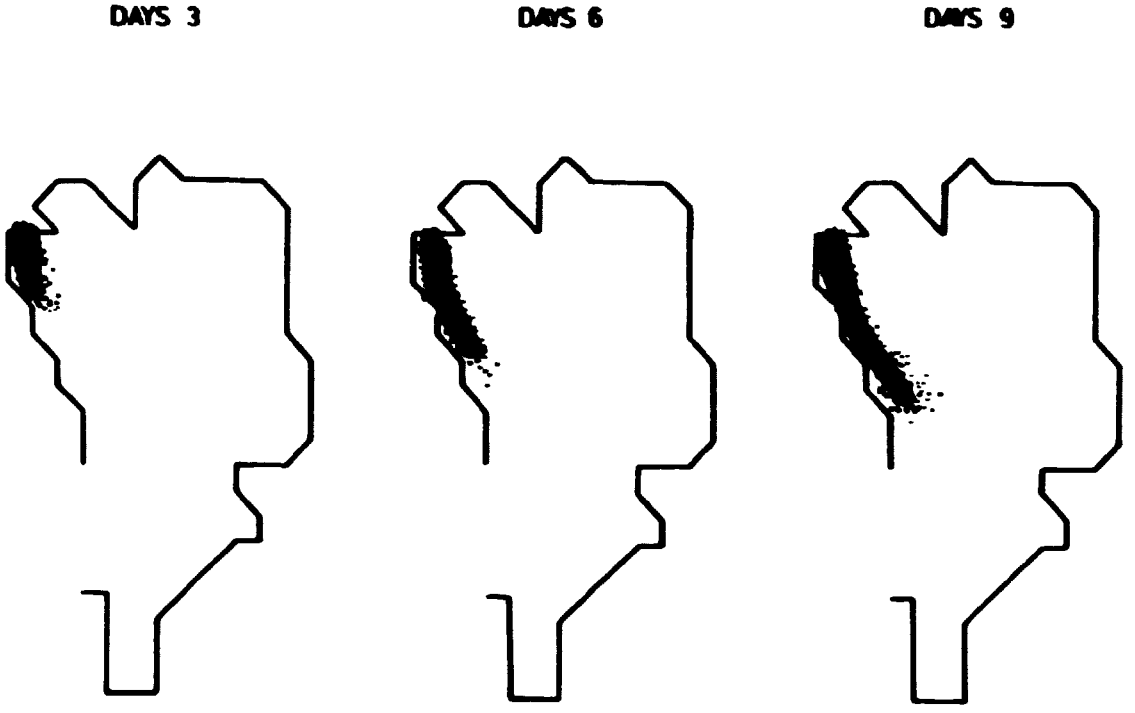


Fig. 3.4. Simulation of particle movements in a Swedish lake (Bork, 1977).

where A must be found from earlier experiences and ΔT is the time interval. A particle is released each 10 minutes and A is equal to 10^5 cm^2/s giving P_u = 32 cm/s. The characteristic time and length scale for this application is 10 days and 40-50 km. A case study involving two outlets to the lake was simulated and concentration isolines were then drawn (Fig. 3.4).

This represents a simple structure of the stochastic element involved in turbulence. Perhaps it leads to reliable concentration estimates, but it assumes that there is no correlation between the turbulent movements of two particles and no Lagrangian (Chapter 2.1) correlation for each single particle. Therefore, assuming that P_u is a reliable estimate, the model will generally overestimate the dilution mechanics. The basic principles for the model seem to be convenient, if what is needed is an "easy to handle" model, and if a general picture of the current field is

available together with the diffusion coefficient. Since turbulence is stochastic in its nature, in general it seems to be reasonable to use some kind of simulation techniques.

3.4.2. Nielsen model

Another approach, applying similar techniques, is described by Nielsen (1983). Here the particles were released every 40 seconds. Nielsen focussed on situations where the current direction changes, since that is the time when local concentrations can increase.

As in earlier cases this model is not designed to handle radioactive discharges, only normal sewage. As input the model needs: a) an outlet position, b) the discharge variation in time and c) the parameters for current and mixing for the recipient as a function of time and space. This last point can of course be very comprehensive, requiring tracer experiments. The time and length scale limits are (in the given example at least) 8 hours and approximately 2 km.

3.4.3. Palmer model

From the foregoing examples it is obvious that much effort would be saved if the simulations could be based on current measurements alone, i.e. without tracer experiments. This has been tried, for instance, by Palmer (1970). His principle is to estimate the diffusion constant from the equation $D = \underline{u}^2 T_L$, where $\underline{u}^2$ is the variance for u, and T_L is the Lagrangian time scale for the turbulence.

This Lagrangian time scale is made equivalent to the Eulerian defined by

$$T_E = \int_0^\infty f_E(t)dt = \int_0^\infty (U(t)U(t+\tau)/\underline{U}^2)dt \qquad (3.8)$$

which can be found just from measurements at a single point (Sec. 2.1). The current measurements used are hourly mean values, and

the model assumes homogeneous and isotropic turbulence at a distance from 0 to 4 km. At this point, it is appropriate to reiterate Appelquist's (1980) observation that "the development and application of mathematical models ... thus becomes a matter of navigating in the difficult waters between the Scylla of oversimplification and the Charybdis of overcomplexity."

3.5. Other models

3.5.1. Buch model

Buch (1980) writes about a number of dye experiments in the Danish straits. Two equations are used in calculating the horizontal diffusion velocity P:

$$C(r,t) = S/(2\pi H_o(Pt)^2)\exp(-r/Pt) \tag{3.9}$$

$$\sigma_{rc} = 6P^2t^2 \tag{3.10}$$

where

$$\sigma_{rc}^2 = \sigma_x\sigma_y \tag{3.11}$$

is calculated from C(r,t).

H_o is the initial thickness of dye layer, and S is the source term.

From these equations the horizontal diffusion coefficient K_n and the characteristic length scale ℓ_n is found

$$K_n = \frac{\sigma_{rc}^2}{4t}, \tag{3.12}$$

$$\ell_n = 3\sigma_{rc} \tag{3.13}$$

where δ_{rc} as before is given by $\delta_x \delta_y$, where

$$\sigma_x^2 = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 c \, dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C \, dx dy} \tag{3.14}$$

and similarly for σ_y.

Furthermore, the report contains tables for diffusion velocities, diffusion coefficients, longitudinal and transversal variance, horizontal length scale etc., all from comprehensive investigations in the Danish straits. ℓ_n differs from 100 to 1000 m and σ_{rc} differs from 30 m at 5 hours to 650 m for the same time.

The investigations have been performed mainly for theoretical reasons and for a general evaluation of different parameters in the Danish straits, more than as a base for evaluating an actual recipient by forecasting pollution dispersion.

3.5.2. Kullenberg

In Kullenberg (1974) experiments lasting 100 hours and covering 5-10 km are described. It is stated that constant diffusion coefficients are unrealistic, and that the vertical dispersion is very important. The following model for the current field is suggested:

$$u = U_o(z) + u_o(z)\cos(\omega t) \tag{3.15}$$

$$v = V_o + v_o(z)\sin(\omega t) \tag{3.16}$$

where U_o and V_o are mean horizontal velocities and z is the depth. The term V_o is chosen as constant since the coordinate system in the case of a vertical shear can be rotated so that the depth variation of the steady current is fully contained in one component, $U_o(z)$.

With convenient assumptions, this gives

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} + v\frac{\partial c}{\partial y} = K\frac{\partial c^2}{\partial z^2} \qquad (3.17)$$

The solution (i.e. expressions for σ_x^2, σ_y^2 and σ_z^2) is given in the reference (too comprehensive to be quoted here).

3.5.3. Puff model

The study of turbulence is not solely the domain of oceanography; it is of relevance in many other fields, including meteorology. Although some elements in the description of turbulence, such as the density and usually the boundary conditions, differ between air and water, many principles from the methods used in relation to air can also be used regarding water. The following is a brief description of a puff model developed at Risø for the purpose of making forecasts for the spread of an atmospheric discharge from a nuclear power plant, as found in Mikkelsen et al. (1980). The idea is that a single puff is given a velocity from an AR(1)-process (Sec. 2.2): $U(t+\Delta t) = \rho_L U(t) + A(t)$ where ρ_L is the Lagrangian autocorrelation coefficient and A(t) is white noise. The area of one puff increases dependent on the intensity of the turbulence. Single puffs released one after the other with a fixed time interval are correlated two and two. The puff is driven by the actual wind velocity at the point of release, measured as 10 minute mean values (in later investigations a sampling frequency of 20 Hz is tried). The output consists of isoconcentration curves and illustrative plumes up to 7 hours after a release and up to a distance of 25 km (see Fig. 3.5). The results of this model are promising, and a lot of work is being done now to test and develop the model further.

3.5.4. Gryning

Another work done by a meteorologist (Gryning, 1981) will be quoted here, due to its instructive section about dispersion theory. Four types of models are described:

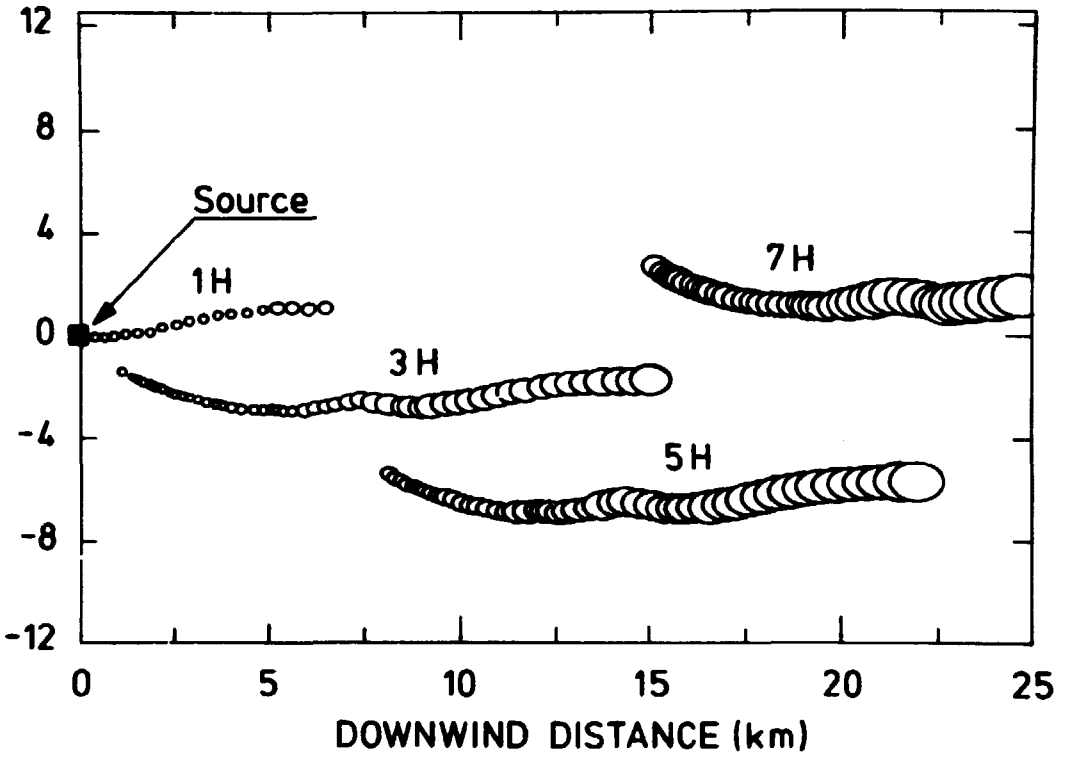


Fig. 3.5. Development of a puff-plume (Mikkelsen et al., 1980).

a) Closure models.

Here the problem is to solve the equation

$$\frac{\partial c}{\partial t} = -u_i \frac{\partial c}{\partial X_i} - \frac{\partial (\overline{U_i c'})}{\partial X_i} + \text{sources} - \text{sinks} \qquad (3.18)$$

The first-order assumption that $U_i c' = -K_{ij}\, \partial C/\partial X_j$ leads to Fickian diffusion, and a solution is given characterised by the Gaussian-shape, i.e. the variances are important parameters. From this it follows that the size of the plume is proportional to $t^{1/2}$. It is stated that K_i cannot be constant, since that would produce eddies of a certain size, in contrast with reality.

b) statistical dispersion theory, founded by Taylor.

Taylor's dispersion formula from 1921

$$\overline{y(T)^2} = \overline{2v'^2} \int_0^T \int_0^t R_{L,v}(\xi) d\xi dt \qquad (3.19)$$

for small $T(R_{L,v} \sim 1)$: $$\overline{Y(T)^2} = \overline{2v'^2}\, T^2 \qquad (3.20)$$

for big T: $$\overline{Y(T)^2} = \overline{2v'^2}\, t_{L,v} T \qquad (3.21)$$

(see also Chapter 2.1)

c) Lagrangian similarity theory pertains to plume height,
d) Convective boundary layer is of no interest here.

After that description there is a section of the "averaging time" in which "absolute dispersion" leading to relative dispersion is examined, where the growth of "a cluster" (different from "a plume") is examined. For relative dispersion (see also Chapter 2) the process can be divided into four areas according to the size of the "dilution effective" eddies:

1) Initial stage, where the particles have not changed the velocity significantly since T = 0. The size of the cloud is proportional to t. 2) Inertial stage, where the dissipation parameter ε is important. The size of the cloud is proportional to $t^{3/2}$. 3) The central stage where bigger eddies are responsible for the dispersion. The size is proportional to t. 4) The final stage, where the velocities of the particles are uncorrelated. The size is proportional to $t^{1/2}$.

Moreover, a relation between Eulerian and Lagrangian description is stated and also some wind variance methods are described where the dispersion is calculated from the variance of the wind velocity.

3.6. Conclusion

We have decribed the type of model best suited to a particular problem, and time and length scales are investigated. One of our conclusions is that it is necessary to decide what scale a model has to cover. In this section we have implied that a model based on current measurements and not on dye experiments should be constructable. A direct comparison of the models is not given; each has its advantages and drawbacks and is, to a certain extent designed for a unique purpose. The very complicated problems one meet when dealing with stratified flow is not reported here. A comprehensive discussion can be found in Bo Pedersen (1980).

4. CORRELATION ANALYSIS OF CURRENT MEASUREMENTS

4.1 Introduction

In this chapter current measurements performed to provide data for the diffusion model will be described. As we know from Section 2.1, the correlation structure of an area is an important property for an analysis of turbulence and is used as input to the diffusion model (Chapter 6). This chapter presents the analysis performed to uncover dependence between distance and correlation. Furthermore, the measurements allow for a statistical time series analysis of the current, i.e. ARIMA-models can be estimated describing the current at a single point. These intensive measurements also provide a detailed knowledge of the current in the area north of the Swedish nuclear power plant Barsebäck, an important area for the initial mixing of effluents from the plant. Many of the radioecological studies in the Danish straits (Chapter 7) are based on releases from Barsebäck.

The correlation structure depends on place, depth, time and weather. It is impossible to find a quantitative correlation function in the literature that is reliable for the area we are

concerned with, although some hints can be given. Moreover, closely placed, highly correlated current meters might provide useless surplus information. On the other hand, this information would be valuable if the degree of surplus, i.e. the correlation, is important to measure. In Boelskifte (1980) correlation analysis is made on current measurements from Faxe Bay (Fig. 4.1), an area south of Copenhagen. However, since the smallest distance between two meters was 5 km, no correlation could be found. This result is confirmed by Riepma (1984) who performed current measurements in the Southern Bight of the North Sea. With a distance of 2 km between current meters he found no correlation at supertidal frequencies. As a preliminary investigation, some data from The Belt Project (Kruse et. al. 1980) have been analysed, where the distance between two meters located near Helsingør and Helsingborg was about 1 km. The original data revealed a correlation of about 0.6, but after a highpass filtration using $f = (12h)^{-1}$ as a cutoff frequency (see later in this chapter) no cross correlation remained. Thus, it became obvious that a new measuring programme had to be set up.

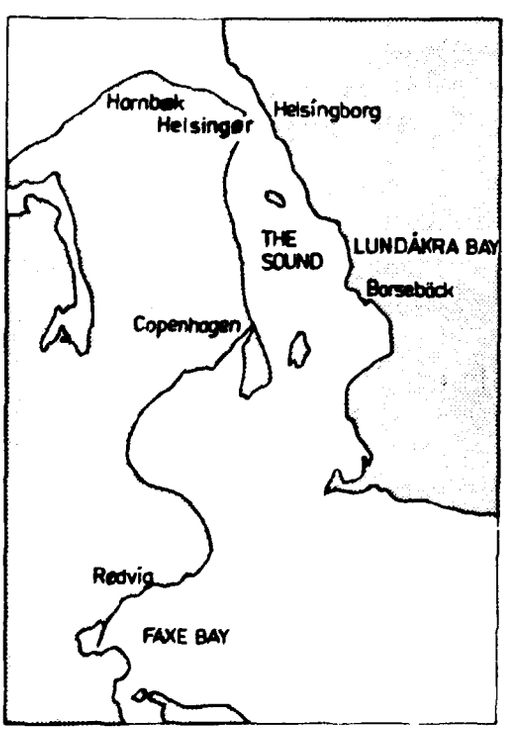


Fig. 4.1. Map showing Barsebäck and its marine environments. The current meters were placed in the Lundåkra Bay.

Section 2 describes the place, time, and principles for the current measurements, together with a short general explanation of the current in the Sound, and some examples of the actual measurements. In Section 3 the efforts to find a reliable correlation structure for the area are outlined. Especially the effect of different kinds of filtering is examined. Section 4 contains a discussion of the effect on the spectrum of the filtration methods.

4.2. Current measurements

4.2.1. Relative positions of the meters

The main current in the Sound changes only between the north and south directions (see below), which means that the correlation along a north-south line will be higher than that along an east-west line. Due to this well-defined main current direction (depending only on the sign), it was decided to have a N/S-line as the longitudinal line for the meters. It was expected that an exponential-like correlation curve would arise to describe the distance dependence. With the main current direction equal to the correlation line between the meters it was presumed that the 900 m maximum distance would discover some correlation, at least for some lower frequencies. The Helsingør Helsingborg data mentioned earlier was not collected along a line following the main current. Thus it shows lower correlations than those expected in the present investigation.

First of all, to get as detailed a picture as possible of the area and to test it for homogeneity, as many current meters as possible were needed. Furthermore, since the characteristic length scale was not known beforehand, it was important to have many points on the distance dependent correlation curve. To test the demand of isotropy the meters could not be placed only along a N/S-axis; some should be placed on the E/W-axis. These requirements, together with the limitations in time and number of current meters available, led to the following approach: In the course of two weeks we used five meters capable of providing

three measuring points in each of the two directions. When the meters were moved closer together after the first week, the number of distances which could be covered was doubled. (A schematic representation of the positioning is found in Fig. 4.2.)

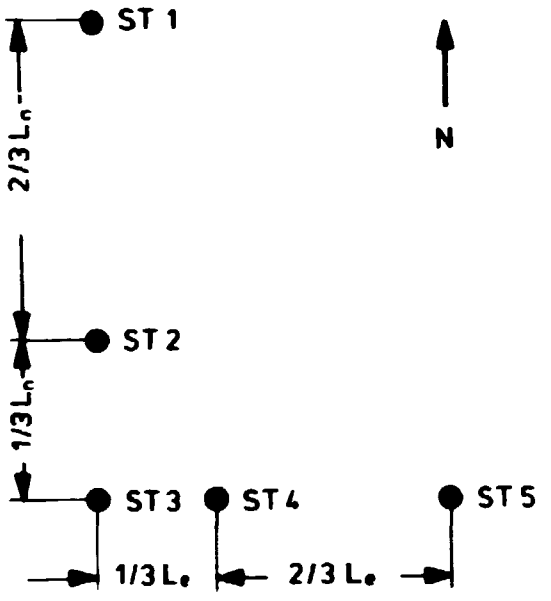


Fig. 4.2. Relative positions of the 5 current meters ST1, ST2, ST3, ST4, and ST5. The distance between ST1 and ST3 is L_n and between ST3 and ST5 it is L_e.

It is seen that using L_n and L_e as maximum lengths in north and east direction, respectively, the distances obtained are $L_n/3$, $2L_n/3$ and L_n for north direction and $L_e/3$, $2L_e/3$ and L_e for the east. During the first week (hereafter called week 1) L_n was, as described, chosen to be 900 m; L_e was chosen as 300 m as the correlation in that direction (transverse to the main current) was expected to be smaller. After week 1 the four meters were moved closer to the one in the middle (Station 3), while the same relative distances were maintained. In week 2 L_n was equal to 300 m and L_e was 90 m.

In order to understand the mechanics leading to the steady-state situation of effluents from Barsebäck in the Sound, as seen in Chapter 7, we must consider long-term changes in the current. This is so, because transport perpendicular to the main current may be the result primarily of transport out of the Sound followed by a change in current direction. But these changes in current direction are unimportant for turbulent diffusion. Therefore one week was considered a sufficient time interval for the correlation investigation at each distance. Thus frequencies below 2/week were eliminated.

4.2.2. Measuring technique

The current meters were of the Aanderaa type and were kindly put at our disposal by the Marine Pollution Laboratory, National Agency of Environmental Protection. Each meter had a sampling frequency of 10 minutes, except ST4, which used 5 minutes. The recorded data are 10 minute average current velocities and instantaneous measurements after each sampling interval of current direction, pressure, temperature, and conductivity. The average velocity is measured by the number of rotations of a rotor on top of the meter, and the direction is given in degrees measured by a compass. The method is illustrated in Fig. 4.3.

In the analysis the velocity and direction are transformed into a velocity in the north direction and one in the east. The Figure t is used to indicate the instantaneous nature of the current, but what is used in the analysis is the following, where V (t) means the velocity and T the sampling interval:

$$V_e(t) = \left(\frac{1}{T}\int_0^T A(t)dt\right) \cos\theta_t \quad , \tag{4.1}$$

and

$$V_n(t) = \left(\frac{1}{T}\int_0^T A(t)dt\right) \sin\theta_t \tag{4.2}$$

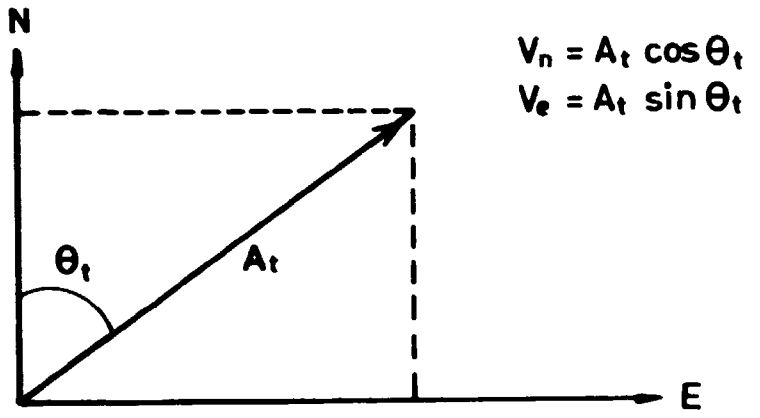


Fig. 4.3. Measurement of current velocity and direction transformed to velocity in north and east direction.

What effect does this sampling method have on the calculation of the correlation? The measuring of the velocity introduces low-pass filtering, and fluctuations during the 10-minute measuring period disappear. In principle, three sampling methods exist: average values for both velocity and direction, an average value for the velocity or direction alone, and only instantaneous values. The first method has the smallest variance, whereas the last has the highest, and higher variance means smaller cross-covariance and crosscorrelation. It cannot be said which of the three methods is the correct one, but in interpreting the results it is important to know how the result was found. A quantitative guess of the difference is difficult to make, but since the direction is relatively constant from one sampling point to the next, it is assumed that the direction is constant during each sampling period. Thus we find a negligible disturbance of the correlation analysis when we use average values also for the direction.

From SMHI (1976) it is known that Lundåkra Bay is an important area for the plume from Barsebäck. While ship traffic is heavy in some parts of the Sound, it was decided to do the measurements as shown in Fig. 4.1. The meter in the middle (ST3) was placed at

position 55° 46'N 12°54'E and the others were placed in relationship to this, as in Fig. 4.2. The measurements started 26/4 1983 at 12.00. The meters were moved 3/5 between 12.00 and 15.00 and were then stopped on 9/5 12.00.

4.2.3. The current in the Sound

The current in the Sound is mainly wind induced. A description of this dependence is given in SMHI (1976). Wind from the north results in high water levels in the southern part of the Baltic producing a north going current through the Sound. Wind from the east also results in high water levels in the southern Baltic, and low levels in Skagerak also lead to a north-going current in the Sound. The result of wind from south is the same, but westerly wind leads to south-going current in the Sound, since it produces high water levels in Skagerak and Kattegat and low levels in the southern part of the Baltic.

This description is valid for the upper layer of the Sound, but the area is strongly stratified with light, low-saline water from the Baltic at the surface, and heavier, high-saline water coming from the North Sea at the bottom.

These two layers can act almost independently of each other. For instance, a situation can occur with northgoing current for the surface layer and southgoing for the bottom layer. The halocline, i.e. the dividing line between the two layers, is very stable during most weather situations. However, when the wind is strong from the west, water enters from the North Sea and changes the salinity. In the beginning the change will be in the form of a thicker bottom layer. This is a very brief description of the complicated current system. A more comprehensive one can be found in Voipio (1981), which Fig. 4.4 is taken from, showing the water exchange between Kattegat and the Baltic, and the different salinities involved.

Another property related to the water exchange between the Baltic and the North Sea is important in relation to dose calculations (Chapter 7), viz. the mean residence time, i.e. the average time

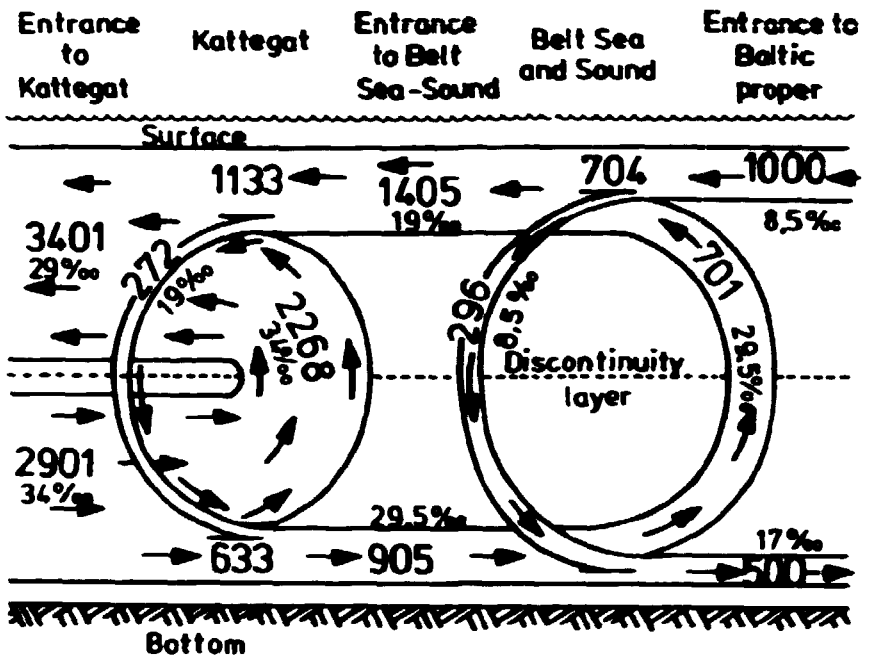


Fig. 4.4. Water exchange through the Danish straits. (Voipio, 1981). Unit: $km^3\ Y^{-1}$.

a volume of water remains in an area. In Clark et al. (1980) the figure for the North Sea is 2.5 years and for the Baltic 25 years. The low value for the North Sea is due to the rapid water exchange with the Atlantic Ocean.

4.2.4. Results

An example of the current measurements from ST1 is shown in Fig. 4.5, where two equivalent representations are given, one for the velocity components in the northerly and easterly directions, and one for the absolute velocity and the direction. All the meters functioned properly during the whole period, with a single exception at the end of week 1 (ST4). If the plots are seen together (not shown here), it is found that all stations show similar pictures during weeks 1 and 2. However, none of them are quite alike, indicating that the distances chosen have neither been too long nor too short, since correlations less than one but

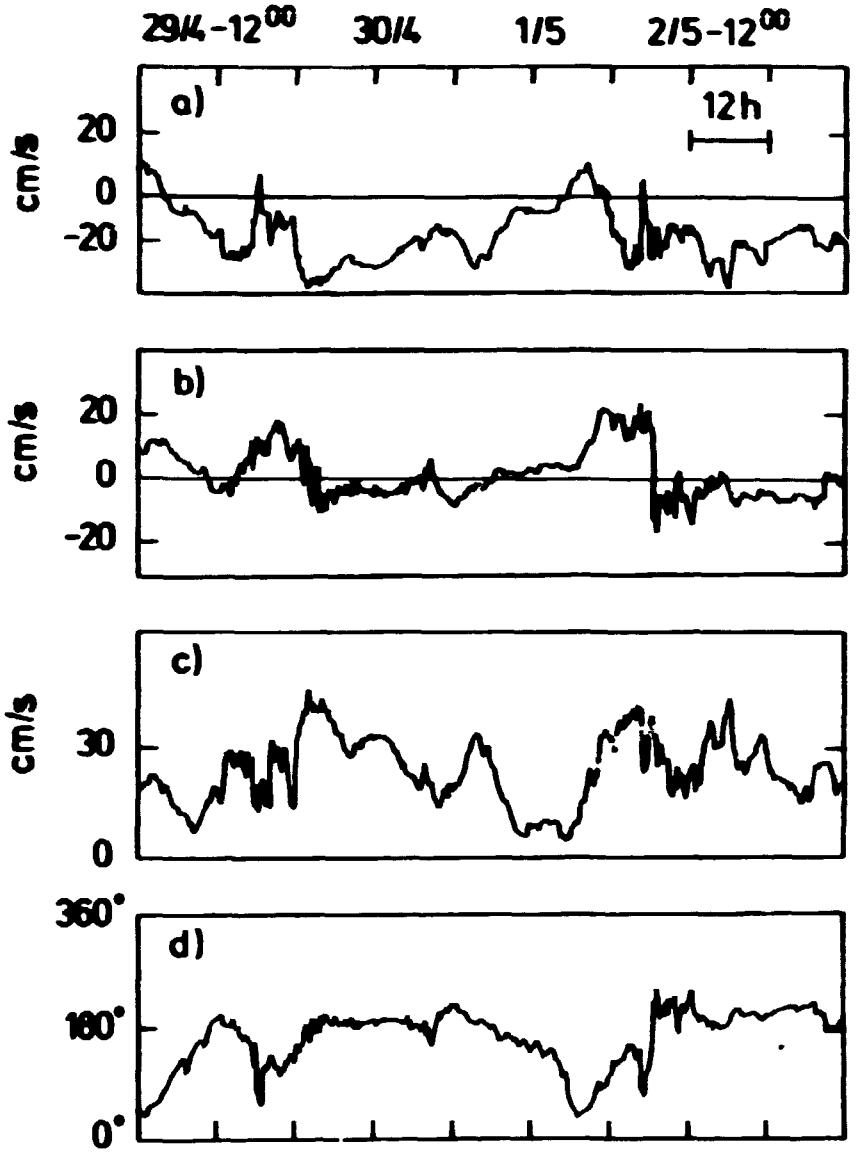


Fig. 4.5. Current for ST1, week 1. a) North component, b) east component, c) velocity, and d) direction.

significantly higher than zero are expected to be found. At all stations the current direction is usually northward (180° on the figure means that the current comes from the south) and, for a short period of time, southward, showing that the situation at the actual point in the Lundåkra Bay is the same as the expected main current situation in the Sound.

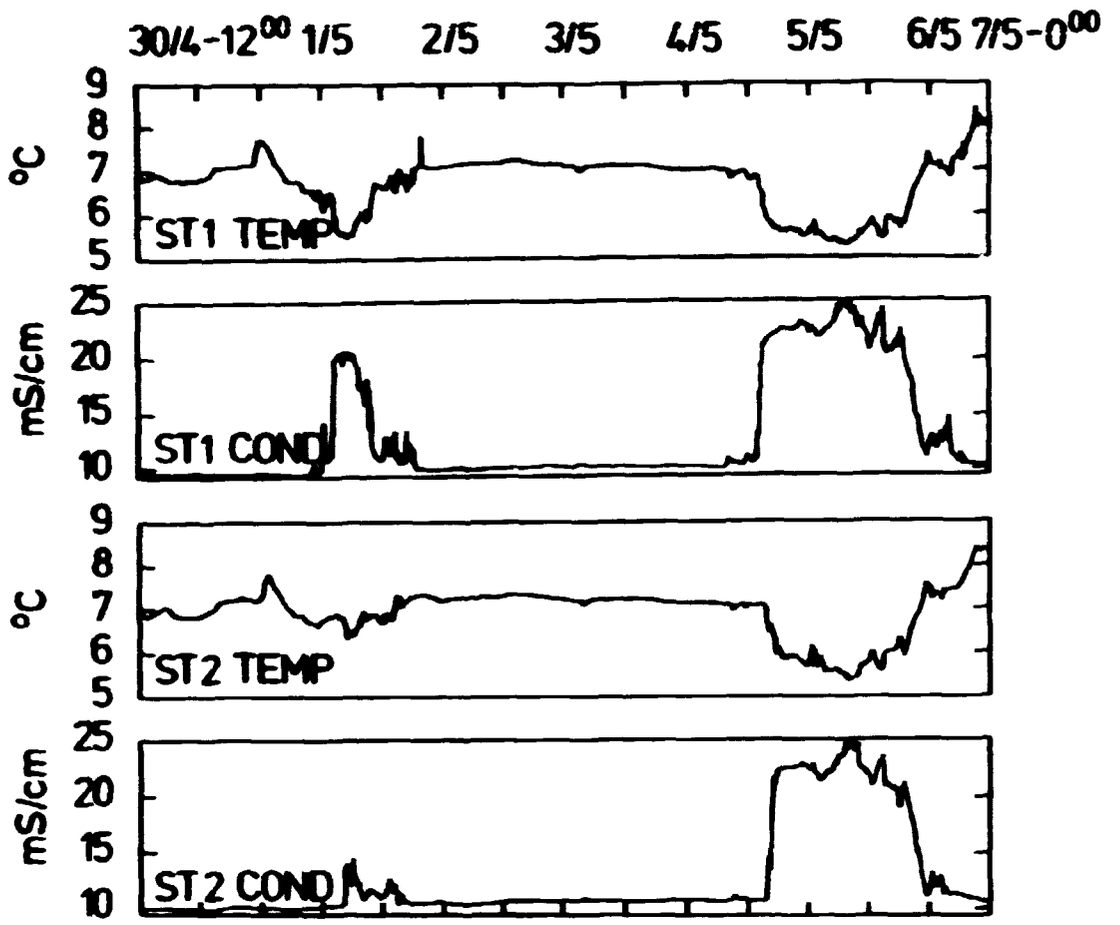


Fig. 4.6. Temperature and conductivity for ST1 and ST2. week 1.

The change in the thickness of the surface and bottom layers during a period with westerly wind means that the meters should be placed at approximately the same depth. For the actual measurements the depths differed between 6.5 and 9.5 m, but that was first seen on the recorded tapes after the measuring period had stopped. In the following we discuss how much these differences can influence the correlation analysis.

By checking the graphs for temperature and conductivity (an example is shown in Fig. 4.6) one can determine whether some of the meters were below the halocline. For all five stations a re-

markable course can be seen. During a long period the conductivity, which is related to the salinity in a complex but proportional way, remains constant at a low level, before suddenly increasing by a factor of three, fluctuating for some time, and finally coming back to its initial value. A similar course is seen in the temperature graphs, which show low levels when salinity is high. In accordance with this, the current direction is southernly at the time when we observe low temperature and high salinity. (It is not a rule, but usually the case, that North Sea water is cooler than Baltic Sea water). This is the characteristic situation for all five meters, but the time periods with high conductivity differs in length and levels, thus revealing that most of the time the meters must be in the same layer, but during some periods they are in different layers.

Another way of using conductivity, temperature and current direction as an indication of different depths is to look at their variances, as shown in Table 4.1. The mean temperature for ST3 is the lowest for all meters and the mean conductivity the highest, indicating that this station is the one most often below the halocline. This is confirmed by the variance in conductivity, which at ST3 is remarkably higher than for the other stations.

Table 4.1. Mean values and variances for temperature and conductivity for all stations, week 1.

	Temperature (^{o}C)		Conductivity (mS/cm)	
	mean	variance	mean	variance
ST1	6.8	0.38	9.9	3.4
ST2	7.0	0.60	9.7	1.2
ST3	6.7	0.94	13.2	5.4
ST4	7.2	0.88	9.3	1.3
ST5	7.4	1.00	9.6	1.5

The picture outlined here for the current situation obviously shows the dynamic of the two layered system in the Sound. It also shows the need for a vertical current profile, which was not possible to measure under the actual practical limitations. The need for similar depths (better seen after the correlation analysis) is another important result of this measuring programme.

4.3. Correlation analysis

4.3.1. Analysis of the raw data

Figure 2.3 illustrates the problem of dividing the current into two parts, one responsible for the transport of the cloud and the other for the diffusion. Very small eddies will not affect the size of the cloud and do not contribute to its transport. Eddies much larger than the cloud also do not affect its size, but act as a transportation mechanism. Eddies with a size comparable to the cloud are significant in the diffusion process and do not contribute to the transport. In reality, it is difficult to separate these different eddies. Where to draw the line between small and large eddies is a matter of subjective evaluation. In other words, one must decide how to define the mean velocity on the one hand and the fluctuations on the other. One man's average is the other man's turbulence.

All original series have been tested for trend, since this can disturb the correlation and the spectral analysis. Only an insignificant linear trend was observed and therefore the trend was not taken into account.

Two properties in the measuring procedure resulted in the first filtration. The sampling interval of 10 minutes means that frequencies higher than $(20\ \text{min})^{-1}$ are impossible to detect. At the other end of the spectrum the measuring period of approximately one week leads to the lowest detectable frequency, (i.e. 1/week). To get an idea of the actual length scale, it can be seen that a typical value for the current velocity is 10 cm/sec leading to 60 km/week.

In Chapter 2 it was noted that for the current in a given direction, e.g. for the northerly component, the cross correlation between the velocity at two specially separated points X and Y is defined as

$$\rho_{x,y}(s) = \frac{E\{(x(t)-\bar{x})(y(t-s)-\bar{y})\}}{\sigma_x \sigma_y} \qquad (4.3)$$

where an overbar indicates average and σ^2 is the variance. Also, as described in Chapter 2, there are two kinds of crosscorrelation, the longitudinal (f) and the transversal (g). When the meters are placed along the north/south axis (ST1, ST2, ST3), the north components determine the longitudinal correlation and the east component the transversal. For the east direction (ST3, ST4, ST5) it is opposite, which means that at a first examination only the combinations listed in Table 4.2 are of interest. If other combinations, such as ST1-ST5, should be used, they can be found by a transformation of the N/E coordinate system to one appropriate for the two points considered. However, for the first analysis the combinations mentioned are sufficient. Weeks 1 and 2 are of course analysed separately. It also has to be mentioned that in all cases the figures given are the correlation when time lag equals zero, i.e. s = 0 in Equation (4.3). Several tests have been made to see whether the maximum of the correlation as a function of time lies at s = 0. This is not always the case (an example of the time dependence is given in Fig. 4.7) but the value for s = 0 does not in any case differ significantly from the maximum value. This is not surprising, since the meters are placed relatively close together in relation to the current velocity.

Before the filtering technique is described, the correlation analysis on the raw data will be examined. For ST1, ST2 and ST3 some high correlations are found, especially for the longitudinal component. These high values may primarily be due to some low-frequency signals, i.e. large eddies producing movements influencing all five meters. Neither f nor g show a clear picture of decreasing values with increasing distance. It is even impossible to find a significant difference between weeks 1 and 2, but f

and g are clearly different, with g having the smallest values, in accordance with theory. For ST3, ST4 and ST5 a difference can be found between the two weeks, but here again the decrease with increased distance is impossible to see. Rather, it seems that correlations where ST3 is involved are lower than the others, even though the distance is smaller. This may be due to the deeper position of ST3 as discussed earlier, but some filtration must take place before more reliable conclusions can be made.

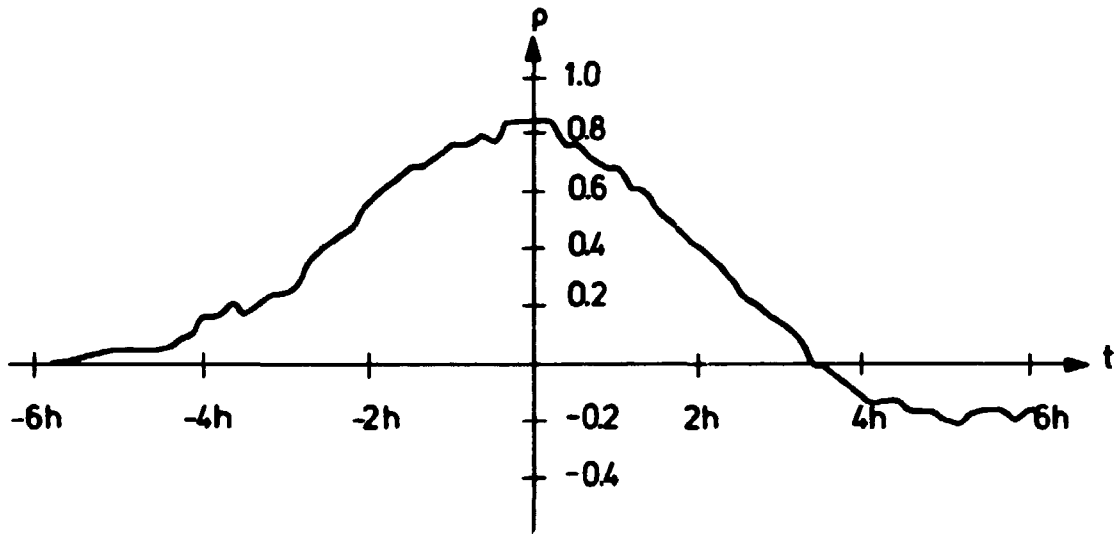


Fig. 4.7. Crosscorrelation as function of time.

4.3.2. Examples of filtering

From earlier sections it is remembered that the lowest frequencies do not contribute to diffusion and should be filtered out. The problem is twofold: 1) At what frequency does the signal have to be cut off? and 2) What kind of filter should be used? In the Sound there is no tide so this component is not found in the spectrum; otherwise it would have to be removed.

A discussion of filtering technique and different filters can be found in Koopmanns (1974). Without discussing the quality of different filters here, we found that a Hanning filter works well. This filter has filter weights

$$a(k) = \frac{1}{m+1} \cdot \frac{1}{2} + \frac{1}{2} \cos(\pi k/(m+1)). \qquad (4.4)$$

where m is the length of the filter. The filter has the spectral window

$$A(f) = m \frac{\sin(2\pi fm)}{2\pi fm} \cdot \frac{1}{1-4f^2m^2} \qquad (4.5)$$

The weight function and the spectral window are shown in Fig. 4.8.

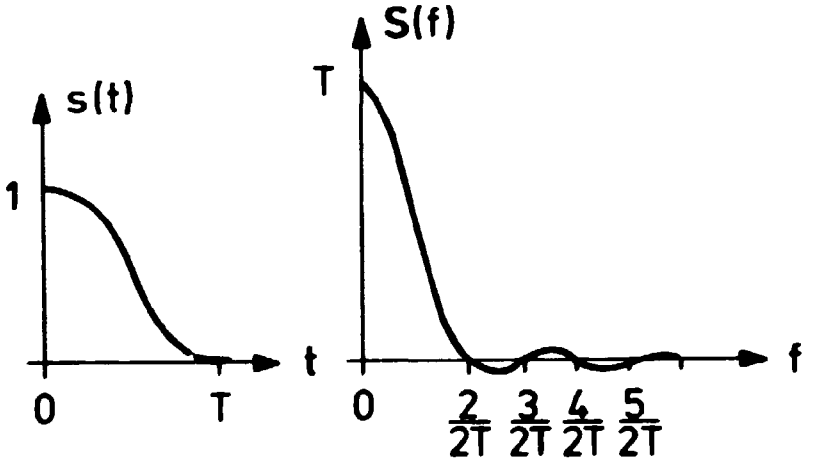


Fig. 4.8. Weight function and spectral window for the Hanning filter.

Many filtrations have been made with the Hanning filter for different values of window length, m. The values of the correlations are seen in Tabel 4.2.

An example of a part of the original series, the lowpass and the highpass is shown in Fig. 4.9, where the highpass is found simply

Table 4.2. Correlations before and after filtration. 1 - 3 N means the correlation in the northern components between stations 1 and 3.

Combinations		cut-off frequency 0	$(33.5h)^{-1}$	$(3^{\wedge}h)^{-1}$	$(27h)^{-1}$	$(24h)^{-1}$	$(20h)^{-1}$	$(17h)^{-1}$	$(13.5h)^{-1}$	$(10h)^{-1}$	$(8.5h)^{-1}$	$(7h)^{-1}$	$(5h)^{-1}$	$(3.5h)^{-1}$	$(1.8h)^{-1}$	$(0.5h)^{-1}$
	1-3 N	.79	.42	.39	.35	.31	.27	.23	.18	.16	.16	.18	.20	.19	.12	.08
	1-2 N	.73	.47	.44	.41	.36	.31	.26	.20	.14	.12	.10	.08	.09	.13	.13
	2-3 N	.76	.55	.52	.49	.45	.42	.39	.35	.31	.28	.23	.18	.10	-.01	-.20
	1-3 E	.24	-.01	-.01	-.01	-.01	-.02	-.05	-.08	-.10	-.10	-.09	-.06	-.06	-.07	-.01
w	1-2 E	.60	.27	.26	.25	.23	.21	.18	.14	.08	.04	.00	.03	.07	.02	.02
e	2-3 E	.43	.21	.20	.19	.16	.12	.07	.01	-.03	-.02	-.01	-.01	-.02	-.03	-.03
e																
k	3-5 N	.29	.15	.13	.11	.09	.05	.01	-.03	-.05	-.05	-.05	-.03	-.02	-.02	.04
	4-5 N	.70	.57	.55	.53	.50	.47	.42	.36	.28	.24	.18	.10	-.01	-.12	-.07
1	3-4 N	.42	.19	.17	.15	.13	.11	.08	.04	.01	.00	.00	.01	.03	.05	.07
	3-5 E	.28	.03	.02	.02	.01	.00	-.01	-.01	.01	-.00	.01	.02	.05	.04	.03
	4-5 E	.61	.32	.30	.28	.25	.23	.21	.19	.16	.12	.05	-.03	-.07	-.05	-.03
	3-4 E	.37	.14	.12	.10	.07	.03	-.02	-.06	-.08	-.07	-.05	-.03	.00	.01	-.01
w	1-3 N	.79	.46					.37			.34			.24	.09	
e	1-2 N	.79	.53					.41			.37			.17	.00	
e	2-3 N	.62	.41					.32			.29			.14	.05	
k	1-3 E	.74	.49					.44			.41			.22	.09	
	1-2 E	.59	.41					.32			.21			.08	-.01	
2	2-3 E	.49	.24					.20			.17			.13	.06	
	3-5 N	.80	.48					.38			.36			.26	.20	
	4-5 N	.96	.85					.79			.69			.45	.26	
	3-4 N	.77	.47					.36			.34			.15	.05	
	3-5 E	.50	.24					.26			.29			.22	.14	
	4-5 E	.86	.80					.74			.56			.22	-.03	
	3-4 E	.39	.19					.21			.22			.14	.04	

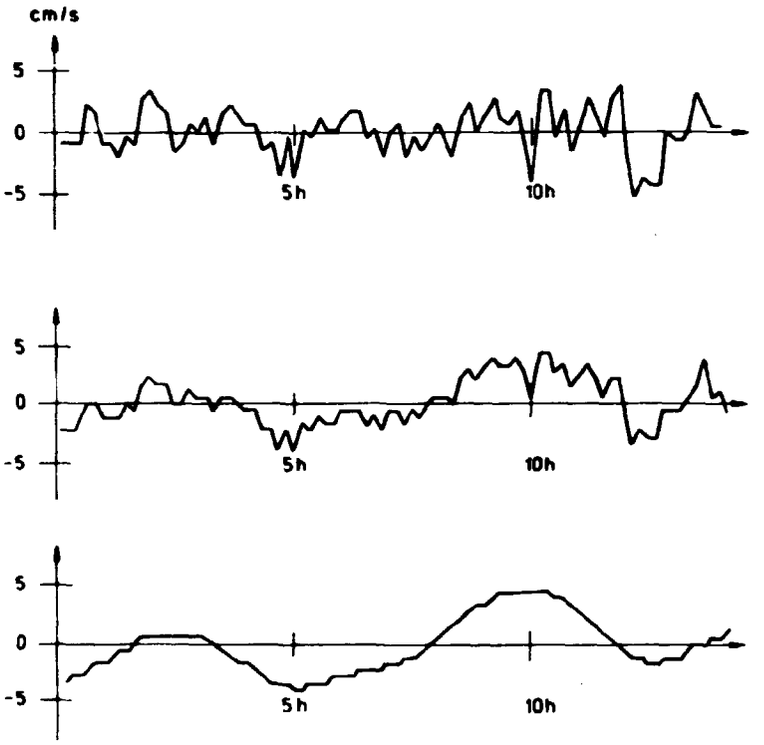


Fig. 4.9. Highpass, original, and lowpass data for ST1 week 1, north component.

as the original minus the lowpass series. The effect of filtering is easily discerned.

For each pair of stations it is obvious that the smaller the value of m, the larger the fraction of the low frequencies that are filtered away, and the smaller is the correlation. This is in accordance with expectations, but raises the question: At what frequency is it relevant to cut off? There is a dramatic drop in the correlation already using m = 100, which has a cut-off frequency of

$$\frac{1}{2\times 100+1} \cdot \frac{1}{10 \text{ min}} \sim \frac{1}{33.5 \text{ h}} \qquad (4.6)$$

i.e. only fluctuations with periods less than 33.5 h will be detected, showing that much of the correlation is due to very slow fluctuations.

In the original data there was a significant difference between the variances of the signals, especially for week 1 where the variance for ST3 was higher than for the other stations. This tendency is less pronounced after a highpass filtering, thus indicating that the difference in depths is most important in the low-frequency part. For week 2 there is no significant difference between the stations, and the magnitude of the variances is the same as for week 1.

In all filtrations the same tendency can be seen: No distinct decrease in the correlation for ST1, ST2, and ST3 or for ST3, ST4, and ST5. If corresponding points are compared for weeks 1 and 2 a small increase in correlation for week 2 can be seen for most cases. For week 1 the longitudinal correlation, f, is for all cases higher than the transverse, g, and for week 2 f is larger than g for most cases. For ST1, ST2, and ST3 f is positive for $m > 10$, but for $m = 5$ some correlations are negative; whereas for ST3, ST4, ST5 some g correlations are negative for $m = 50$.

Although the depths differ a little, the lack of consistency is surprising since the time series did not show up as being very different. It is also surprising that after the first filtration ($m = 100$) some of the correlations are already almost zero (one value is even negative). In all cases the lowest values occur where ST3 is involved.

4.3.4. Other filtering methods

4.3.4.1. Shorter periods: To eliminate possible instationarities, it would be worthwhile to perform the correlation analysis for smaller periods of time during which the main current does not change direction and the intensity of turbulence is smaller than it is during periods of unstable current direction. The periods can be found by examining the current direction. However, conductivity and temperature can also help in finding relevant periods. From week 1 two periods were found, while for week 2 only one was found; all periods were of low and constant conductivity, and a current direction that did not alter too much. The length of each

series was about 24 h, comprising 144 observations. Periods from the same week made it possible to check if the correlation was constant in time. (Table 4.3.)

Table 4.3. Correlation on shorter periods.

Combination	Week 1 28/4-12^{00} 29/4-12^{00}	Week 1 2/5-01^{20} 2/5-23^{00}	Week 2 3/5-20^{00} 4/5-11^{00}
1-3 N	.76	.82	.92
1-2 N	.86	.95	.98
2-3 N	.75	.88	.92
1-3 E	-.29	.68	.59
1-2 E	.75	.83	.80
2-3 E	.07	.85	.54
3-5 N	.37	.27	.86
4-5 N	.47	.61	.93
3-4 N	.67	.64	.88
3-5 E	.12	.30	.40
4-5 E	.50	.48	.58
3-4 E	.61	.56	.52

The longitudinal correlation is still higher than the transverse, and for some combinations, especially in week 2, the correlations are now higher than for the long series. This means that if the meters are placed in the same layer, which is the case for the present periods, they are highly correlated and this correlation is not due to very slow variations in the mean current. A few of the figures from the week 1 periods differ very significantly but most figure are more or less equal for the two periods. As usual, f is larger than g and the increase from week 1 to week 2 is obvious, at least for the f values, indicating the dependence on the distance. For ST3, ST4, ST5 an increase in correlation is

found in this way. Although it still is far from an ideal picture, it shows that it is a better approach than using the whole series.

4.3.4.2. Mean lowpass subtracted. Another method has been tried in finding a correlation scheme. Referring to the concept of homogeneity it is assumed that the mean current is the same for all five stations. Thus more information about this can be obtained. The mean of the five meters has been found and subtracted from all series. This eliminated all correlation and left only uncorrelated white noise. A filtration had to be done, so the following procedure was used: First, all ten series (two from each station) were divided into a low- and highpass part:

$$X_i = X_i^{\ell} + X_i^{h} \quad , \quad i = 1,2 \ldots\ldots 10 \tag{4.7}$$

A Hanning filter as described earlier was used. Now an average was created from the lowpass series of each of the two groups (north and east components) of five series:

$$\bar{X}_{\ell} = 1/5(X_1^{\ell} + X_2^{\ell} + X_3^{\ell} + X_4^{\ell} + X_5^{\ell}) \tag{4.8}$$

equivalent for the northerly and easterly directions. With a sensible choice of the highpass filtering length, m, $\bar{X}_{\ell}$ is the best estimate of the mean current, including information from all five stations and without highpass data. Of course, it makes no difference whether the lowpass filtering or the averaging is done first. This mean current is subtracted from all series

$$\tilde{X}_i = X_i - \bar{X}^{\ell} = X_i^{h} + X_i^{\ell} - \bar{X}^{\ell} \tag{4.9}$$

where $\bar{X}^{\ell}$ can be called the deterministic part of the current, equal for all stations, and $\tilde{X}_i$ is the stochastic part on which correlation analysis and time series estimation have to be made.

For some of the combinations negative correlations are observed in these data. This could be due to a bias introduced by taken the average of all five meters. That would be the result if the meters were grouped systematically around the lowpass mean.

4.3.4.3. Mean lowpass of three stations subtracted. An alternative approach could be to use only three stations to determine the mean current, to compute the correlation between the two other stations when the lowpass of this mean is subtracted, and thus eliminate the bias. This was tried on the whole series for week 1, and then for three shorter periods from the same week. The results are shown in Table 4.4, where only the north components are shown.

This approach also failed to give a distinct picture of the correlation structure. There are pronounced differences between the three shorter periods, which indicates the non-stationarity of the correlation. The decrease for increased distance is again impossible to detect and, in spite of the removed bias, some of the non-biased values are negative. The f values are positive with a few exceptions, whereas most of the g values are negative, with a few having rather high values (.93 and .91). It is impossible to explain some of these discrepancies. Thus, it cannot be concluded whether this approach is useful for other experiments.

The last thing we tried was to use just the current direction as an expression of the mean current. The mean direction for the five stations was found and this was lowpass filtered. Then each station was transformed to a coordinate system changing with the mean velocity, so that it had a component in the direction of the mean direction and a component perpendicular to it. The correlations yielded by this system do not represent f and g, so an interpretation must be provided afterwards. Also, this approach failed to give a clear picture of the correlation. Therefore no results are shown.

Table 4.4. Correlation analysis on manipulated data (see text for explanation).

-	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4
1-3 N	.21	.88	.62	.92	.37	.85	.37	.74	.30*	.86*	.72*	.75*
1-2 N	.33*	.54*	.05*	-.46*	.57	.54	.32	-.26	.55	.55	.39	-.26
2-3 N	.45	-.65	-.20	-.46	.60*	.68*	.52*	-.18*	.57	.67	.55	-.20
3-5 N	-.57	-.22	-.64	-.20	-.30	-.15	-.18	.51	-.34	-.17	-.21	.50
4-5 N	-.57	-.55	-.12	-.69	-.34	-.39	-.09	-.65	-.34	-.46	-.05	-.66
3-4 N	.33	-.22	.93	.15	.11	-.22	.91	-.30	.16	-.30	.79	-.32

A: Average of ST3, ST4 and ST5
B: - - ST1, ST4 and ST5
C: - - ST2, ST4 and ST5
1: week 1 all observations
2: week 1 obs. 236-472
3: week 1 obs. 616-760
4: week 1 obs. 904-998
*: combinations without bias

4.4 Spectral representation

The influence on the correlation of the choice of cut off frequency is seen in Table 4.2; an example of the effect of filtration on the original series can be found in Fig. 4.9. It is also possible to find this influence in the frequency domain. For ST1 Fig. 4.10 shows the spectrum of the original series and the spectra after two filtrations using m = 100 and m = 25, respectively. The higher the cut off frequency the less energy is left in the signal.

Furthermore, comparing the spectra we see that above the cut-off frequency the spectrum is not disturbed significantly by the filtering, thus indicating the usefulness of the Hanning filter.

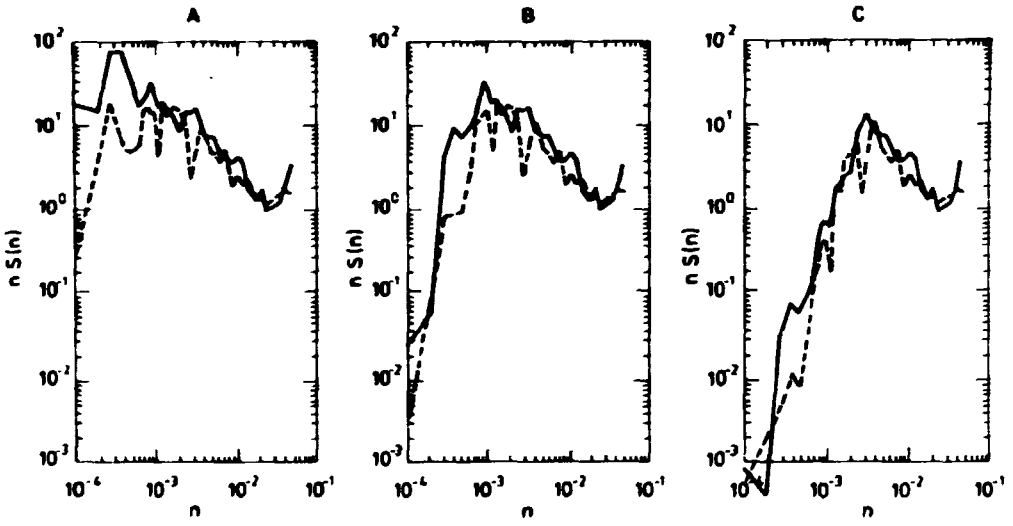


Fig. 4.10. The influence on the spectrum from the choice of cut-off frequency, a) original data, b) m = 100, c) m = 20. Solid line: north component, dotted line: east component.

According to Section 2.1 and Chapter 3 at least two important properties of the turbulence are reflected in the spectrum. In the inertial subrange, where no dissipation and only energy transfer from larger to smaller eddies takes place, the spectrum

should have the form $S(n) = \alpha\varepsilon^{2/3}n^{-5/3}$. Since the spectra are drawn as $nS(n) = f(n)$ it would follow a line proportional to $n^{-2/3}$. This can easily be checked and an example is seen in Fig. 4.11, where the agreement with theory is remarkable.

Also local isotropy should be reflected in the spectrum. According to Tennekes and Lumley (1972) there would be a constant ratio of 4/3 between the spectral estimates for u (current in the direction of the mean current) and v (current perpendicular to this) in the inertial subrange.

If we look at Fig. 4.10 or Fig. 4.11 we see that the ratio is hard to estimate exactly, but 4/3 could be a reliable value.

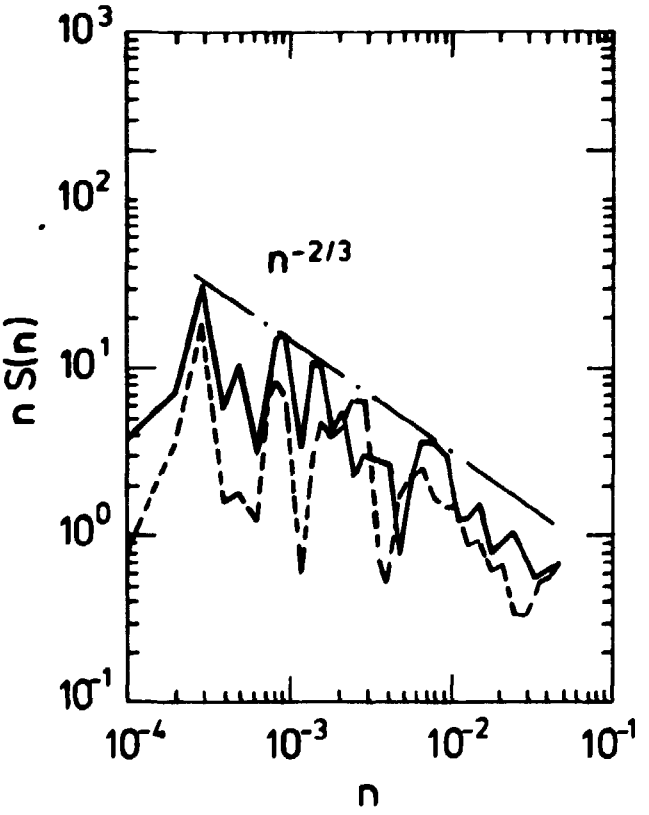


Fig. 4.11. A spectrum showing the Kolmogorov law (Eq. 2.7).

4.5. Conclusion

Although the dependence between distance and correlation did not show up very clearly, some useful information has been deduced from the current measurements. The different correlations between weeks 1 and 2 show that smaller distances mean higher correlations. The correlation analysis has been made on the data after different kinds of manipulation. The data have been highpass filtered using different filter lengths. Shorter periods have been investigated and different expressions of the mean current have been subtractred from the original series. All approaches showed that the correlation is far from an easy to measure, static phenomenon.

The time and length scale used was convenient for the analysis, since both very high and very low correlations were found. This is an important result for future investigations with similar goals.

It is significant that one must place the current meters at the same depth, and that the vertical current profile should also be measured, if possible.

In the chapters that follow a cut-off frequency of 1/7 h is used as a leading example. We assume this to be the most appropriate way of separating the fluctuations from the mean current.

5. A STATISTICAL MODEL FOR THE CURRENT

5.1. Introduction

This chapter is devoted to a statistical description of the current at a single point. First of all the analysis includes an estimation of ARIMA (autoregressive integrated moving average) models on the data, as the basis for building a transport simu-

lation model for the area north to Barsebäck. The number of stations and length of the measuring period allow for many tests for homogeneity, stationarity, etc. ARIMA-model estimation can also be used to investigate the correlation structure of the area, because transfer function models, or multivariate ARIMA-models, can tell whether or not information from one current meter can help explain the current measured by another. So the purpose of the analysis is this: After having investigated the correlation structure of the area (Chapter 4), the transport part should be analysed, correlation structure tested, and models suitable for simulating trajectories should be estimated.

5.2. The data

The ARIMA estimations described in the following section are based on different types of data. Each of the five stations were analysed in detail, but for the sake of simplicity mainly the results from ST1 are shown. Of course an estimate is made using the original data from week 1 and week 2, and on the data after filtrations in which different values for m are used. Furthermore, shorter periods are analysed, and for some of these data multivariate ARMA-models have been estimated. Also the lowpass part has been analysed.

This kind of analysis is in the time domain, but where appropriate, spectra giving the representation in the frequency domain are also used. In the analysis described here we are considering data which are highpass filtered with cut-off frequency $(7h)^{-1}$, i.e. $m = 20$ in the formula for the Hanning window (Eq. 4.4).

5.3. Estimation for high-pass filtered data

5.3.1. Some effects of filtration

As seen in Chapter 4 the current is divided into an average part and a fluctuating part, each of which should be analysed separately. When one uses short periods the mean current is more con-

stant and therefore the data are easier to handle. Whereas, when we examine longer periods, the property for the mean current depends to a great extent on the choice of m.

Before any estimation results are given, it should be repeated from Section 2.2 that using an ARMA-model a time series can be written as such:

$$Z_t = \phi_1 Z_{t-1} + \ldots + \phi_p Z_{t-p} - \theta_1 A_{t-1} - \ldots - \theta_q A_{t-q} + A_t, \tag{5.1}$$

where Z_t is the value of the process (the output) at different moments, A_t is a white noise process, and ϕ_i and θ_i are real numbers. Usually p and q are small numbers, e.g. both equal to one, and in that case the process is called an ARMA (1,1)-process:

$$Z_t = \phi_1 Z_{t-1} - \theta_1 A_{t-1} + A_t \tag{5.2}$$

If Z_t is a multivariate process the equation will be

$$\underline{Z}_t = \underline{\underline{\phi}} \underline{Z}_{t-1} - \underline{\underline{\theta}} \underline{A}_{t-1} + \underline{A}_t \tag{5.3}$$

where _ indicates a vector and = indicates matrices.

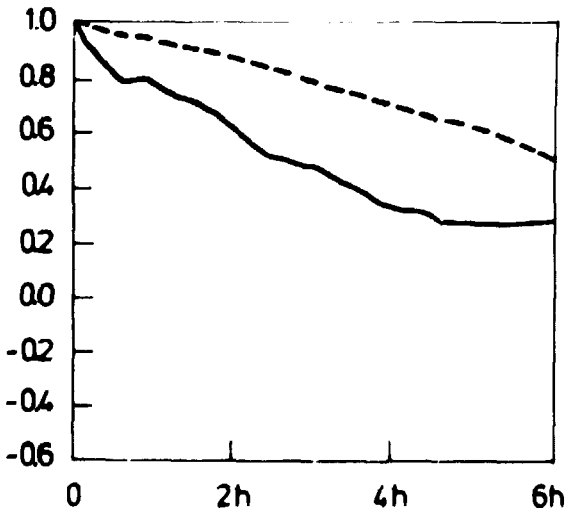


Fig. 5.1. Autocorrelations for ST1, week 1.
---- North component. ——East component.

The multivariate process can express a complex relationship between the elements in the vector $\underline{Z}_t$.

As stated earlier ST1 week 1 is used as the general example in the following. First, it should be repeated from Section 4.4 how the spectrum of the original data looks. The low frequencies indicate that the autocorrelation function (ACF) will reflect the long memory of the signal. This is confirmed by Fig. 5.1 where the ACF for both components are shown, having high values even for large t. The north component has the highest values, because north is the direction for the main current. We have tested whether there is any crosscorrelation between the two components, and the result is shown in Fig. 5.2. It is seen that some correlation exists in the data. The change in this, after the highpass filtration, will be checked later.

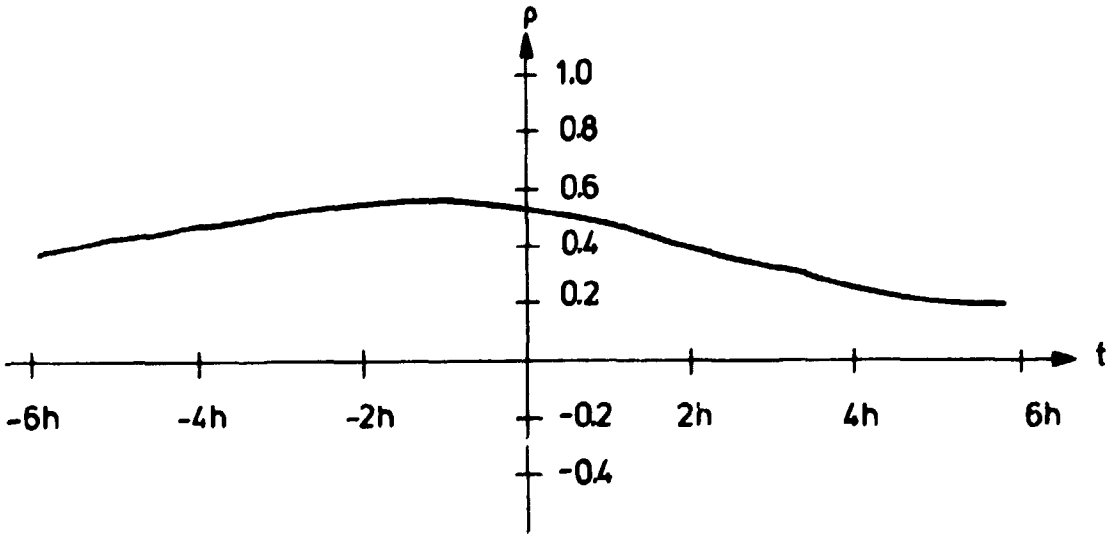


Fig. 5.2. Crosscorrelation between north- and east component for ST1, week 1, original data.

The ACF for the highpass filtered data are shown in Fig. 5.3. The sprectra are already seen in Fig. 4.10. The lower part of the spectrum for the original data has decreased while the rest remains almost unchanged. This phenomenon is also discussed in Chapter 4. For the ACF a remarkable change is seen. Already after

three or four lags the values are equal to zero indicating a process characterised by much more rapid changes, the north component again having the longest memory. As could be expected, the CCF between the north and east component have values which are insignificantly greater than zero for all t. Estimation is simplified when the two components can be estimated independently since off-diagonal elements in the $\underset{=}{\phi}$- and $\underset{=}{\theta}$-matrices are set equal to zero.

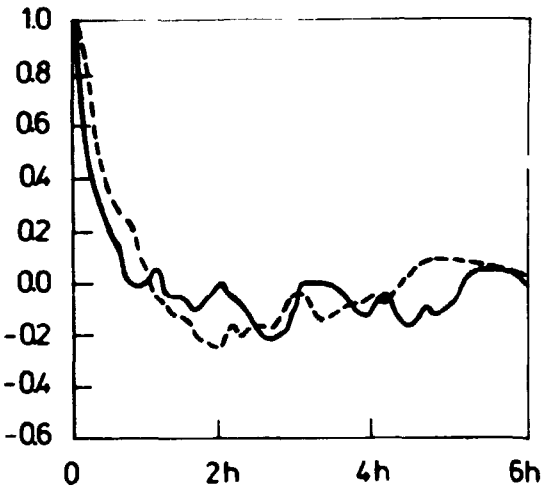


Fig. 5.3. Autocorrelation after highpass filtering for ST1, week 1. ----- north, ——— east.

5.3.2. Identification

For ARIMA-models the first step in the model-building procedure is the identification (Section 2.2). In following the description given earlier we must first find the order of differencing, i.e. subtract neighbour points from each other.

What we are left with after differencing is almost white noise; thus, the ACF's reflect the nature of the process: for t = 0 the ACF (not shown) equals one (per definition) and for t ≠ 0 only small values exist. For 700 observations the variance of the ACF is $1/\sqrt{n} = 0.038$ and therefore values below this limit are assumed to be equal to zero. Thus it can be concluded that a differencing

will result in a series of almost white noise. This can also be seen from the parameter values from an estimation of an ARMA(1,1)-model on the differenced data: If the input series y_t equals white noise A_t then

$$Y_t = A_t \quad \rightarrow$$

$$Y_t - Y_{t-1} = A_t - A_{t-1} \quad \rightarrow \qquad (5.4)$$

$$Z_t = A_t - 1 \cdot A_{t-1} = \text{MA(1) for } \theta = 1 \quad .$$

Thus high values for θ will result. The following model was estimated on the differenced data (the north component):

$$Z_t = 0.47Z_{t-1} - 0.80A_{t-1} + A_t \qquad (5.5)$$

The high value 0.8 confirms that differencing is unnecessary.

The ACF (Fig. 5.3) is an important tool for identification. For u (the north component) the ACF tails off after 5 lags and for v (the east component) after 3 - 4 lags, and both tend towards zero as damped sine waves. The partial autocorrelation function (PACF)

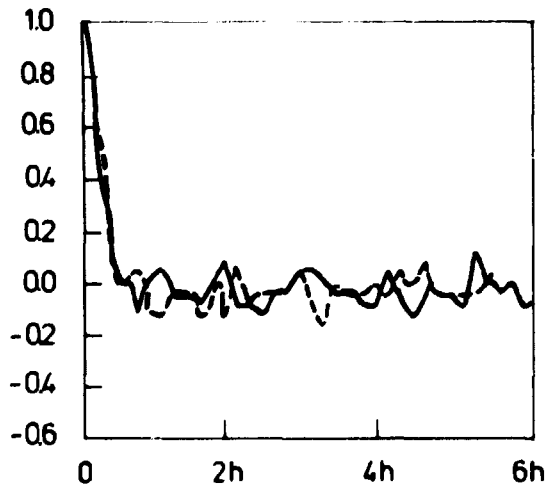


Fig. 5.4. Partial autocorrelation function after highpass filtering for ST1. Week 1. ----- north ——— east

can also be used, as shown in Fig. 5.4. For u only one value is higher than zero, and for v 1 - 2 values are higher. Furthermore, there is no cross correlation between the two series. The ACF-figures show that if it was an ARMA-model q would be lower than p (see Section 2.2), and the PACF-figures indicate AR-models, p being 1 or 2.

The preliminary analysis thus suggests an AR(2)-model. The reliability of this will, of course, be tested by estimating other models as well.

5.3.3. Estimation

The procedure followed in the parameter estimation is, as stated earlier, the one developed by Spliid (1983). A user's guide is found in Spliid (1981), where the identification factor, F, which is referred to below, is described. In the program a partial F-test value for each parameter in the model is computed. Specifying F causes the program to exclude all parameters which have F-values smaller than this.

First, the suggested AR(2)-model was used, and it was found that it fits the data well (diagnostic checking, see later).

For the u and v series (the north and east components) the estimated model is

$$\begin{aligned} &u: \; Z_t = 0.59\, Z_{t-1} + 0.05\, Z_{t-2} + A_t \\ &v: \; Z_t = 0.37\, Z_{t-1} + 0.12\, Z_{t-2} + A_t \end{aligned} \qquad (5.6)$$

using independent estimation for the two series.

This can be written more easily as:

$$\underline{Z}_t = \begin{bmatrix} 0.59 & 0 \\ 0 & 0.37 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} 0.05 & 0 \\ 0 & 0.12 \end{bmatrix} \underline{Z}_{t-1} + \underline{A}_t \qquad (5.7)$$

This notation will be used below.

5.3.4. Diagnostic checking

Having found the AR-model, a lot can be done to test this model and to see how well it describes data. Two approaches are used: one is to analyse the residuals in different ways; the other is to compare properties of the data with theoretical properties of the estimated model. For an AR(2)-model the theoretical ACF and PACF can easily be found in the univariate case (the Yule-Walker equations):

$$\rho_1 = \phi_1 + \phi_2\rho_1$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \qquad (5.8)$$

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + ... + \phi_k$$

The results are shown in Fig. 5.5, and a good approximation is seen, at least for the first lags.

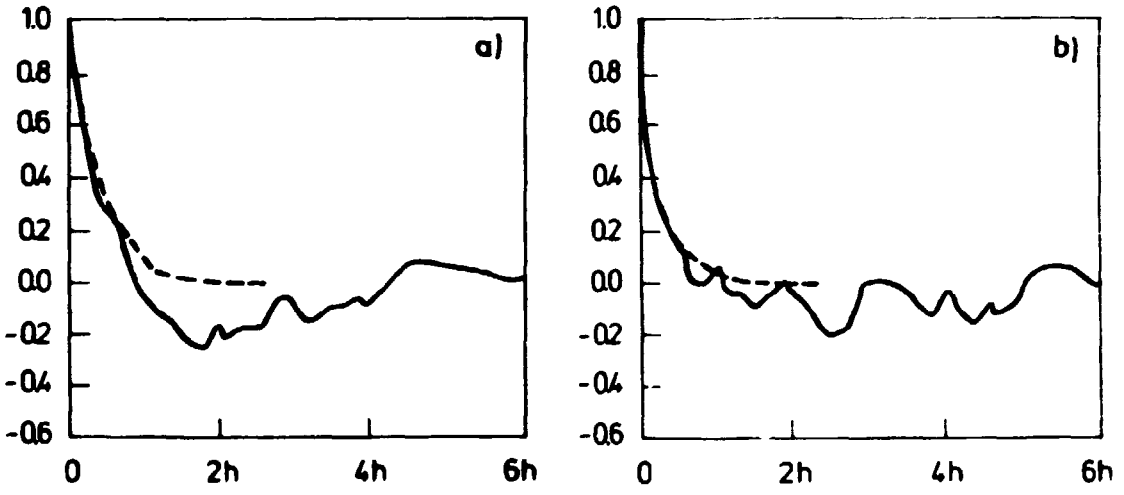


Fig. 5.5. Theoretical (---) and actual (__) autocorrelation. a) north component. b) east component.

The spectrum can also be found:

$$\Gamma_y(f) = \sigma_A^2 \left| \frac{1}{1-\phi_1 e^{-j2\pi f} - \phi_2 e^{-j2\pi f}} \right|^2 \qquad -1/2 < f < 1/2 \qquad (5.9)$$

and the results are shown in Fig. 5.6. The theoretical spectrum is of course very smooth, and some of the differences are due to the way the spectrum of the signal is estimated. Using another filter length in the spectrum estimation would yield a smoother, better fitting spectrum.

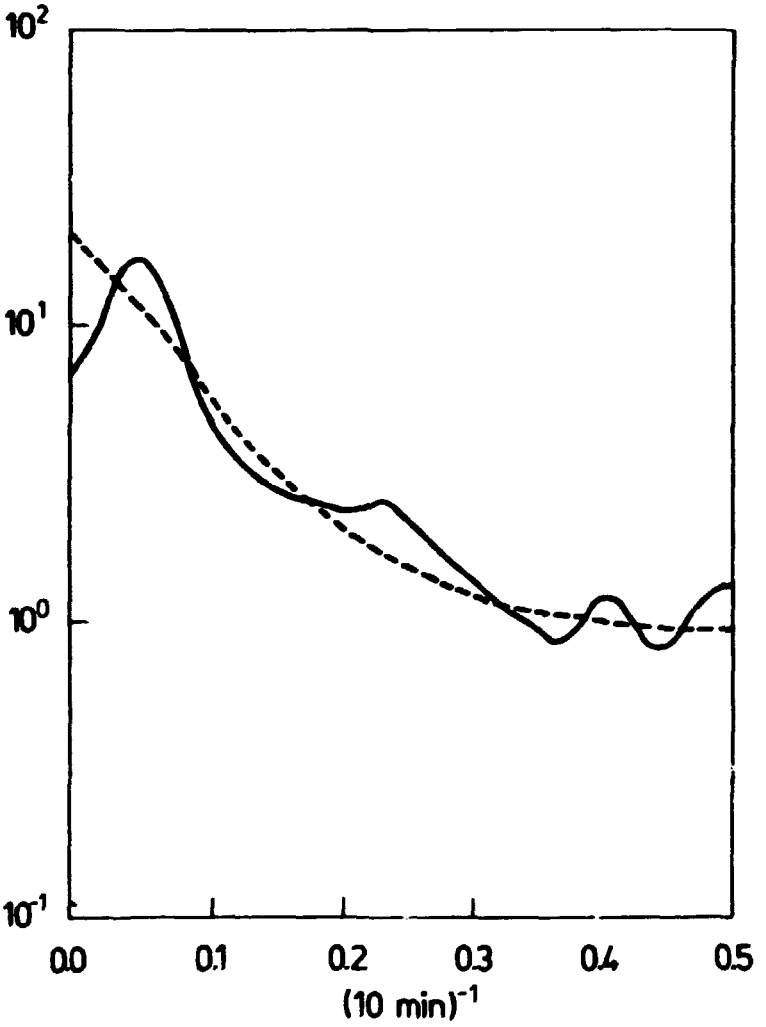


Fig. 5.6. Actual (——) and theoretical (-----) spectrum for the AR(2)-model.

One might be tempted to stop with the very satisfactory results that we have shown here. However, the possibility of checking the model by analysing the residuals should also be utilized. First an AR(6)-model was estimated on the residuals. If the residuals were white (as they should be) the ϕ's in the AR(6)-model would

not be significant (seen by the acoompanying F-value). The result is shown below with the F-values in paranthesis after the estimated ϕ-values.

Table. 5.1. Two-dimensional AR(6)-model estimated for the residents from two univariate models.

$$\begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$$

	ϕ_{ij}	F-value	ϕ_{ij}	F-value
AR(1):	0.0	(0.27)	-0.04	(3.35)
	0.0	(0.07)	0.0	(0.03)
AR(2):	0.0	(0.08)	0.0	(0.02)
	0.0	(0.02)	0.0	(0.32)
AR(3):	0.0	(0.10)	-0.06	(8.60)
	0.0	(0.04)	0.0	(0.74)
AR(4):	0.12	(11.00)	-0.09	(18.31)
	0.0	(0.027)	-0.11	(7.86)
AR(5):	0.0	(0.15)	0.0	(0.06)
	-0.12	(3.09)	0.0	(0.73)
AR(6):	-0.06	(2.36)	0.0	(0.75)
	0.0	(0.53)	0.0	(1.51)

The F-values can be compared to an F(1,500) distribution which is approximately equal to a $\chi^2(1)$ distribution. $F(1,500)_{0.95} \sim 4.0$.

Although some of the F-values are greater than four, the parameter values are low, so no cross correlations can be detected in the residuals. These high F-values could be due to the filtering procedure.

Another interesting quality of the model is the relationship between the residual variance and the variance of the signal. It

was found to be for u: 2.27/3.70 · 100% = 61% and for v: 7.25/8.96 · 100% = 81%. This reduction of variance can easily be found for an AR(1)-process:

$$Z_t = \phi Z_{t-1} + A_t$$

$$(1-\phi B)Z_t = A_t \qquad (5.10)$$

$$Z_t = \frac{A_t}{1-\phi B} = A_t + \phi A_{t-1} + \phi^2 A_{t-2} + \ldots.$$

$$V(Z_t) = V(A_t)(1 + \phi^2 + \phi^4 + \ldots.) \qquad (5.11)$$

$$V(A_t) = V(Z_t)(1-\phi^2)$$

For $\phi = 0.59$ this gives the ratio 0.65 and for $\phi = 0.37$ it gives the ratio 0.86.

As can be derived from later examples, this ratio depends to a high degree on the filtration.

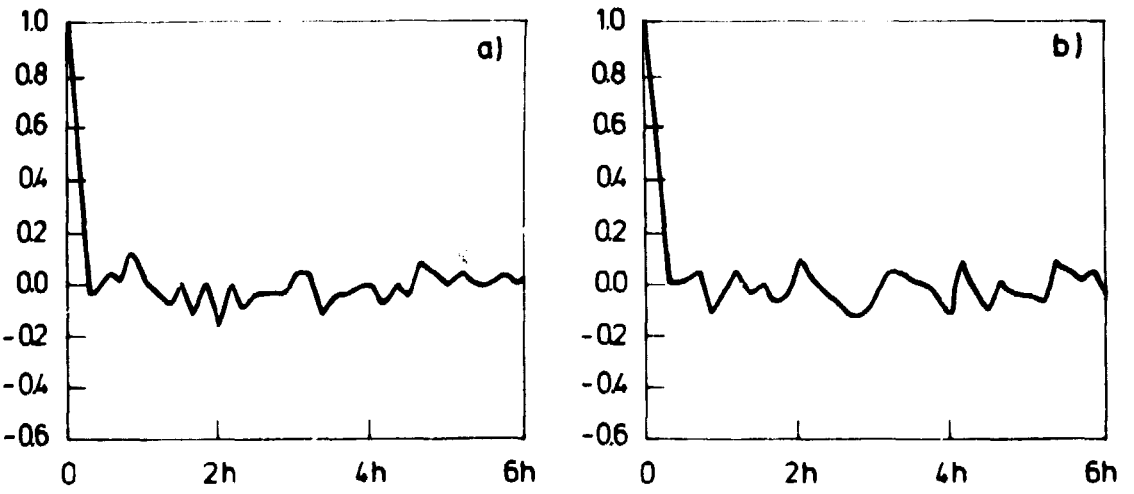


Fig. 5.7. Autocorrelation functions for residuals from the estimated AR(2)-model. a) north component, b) east component.

If we assume that the residuals consist of white noise, their ACF would be zero for lags > 1. The result is shown in Fig. 5.7. The test for white noise is: $\pm 2\sqrt{1/n}$, arising from the N(0,1/n)-distribution, which the autocorrelation estimates approximately will follow. Again "white-ness" of the residuals is confirmed.

In order to be complete we have also included the spectrum of the residuals (Fig. 5.8). It is also characterised by the "white-ness".

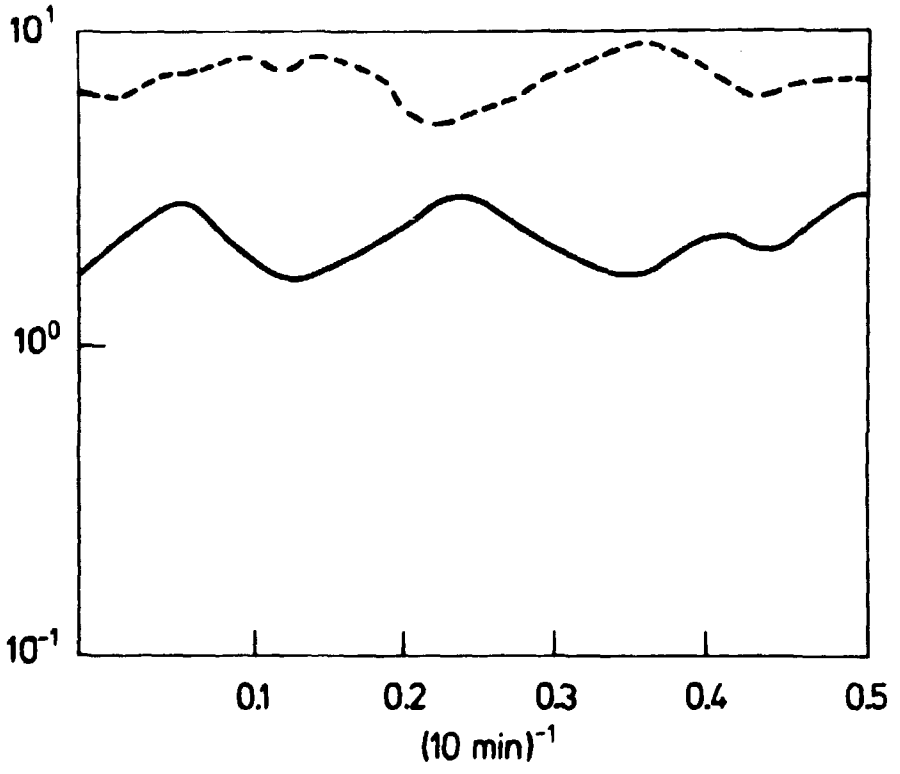


Fig. 5.8. Spectra for residuals from the estimated AR(2)-model ——north component. ----- east component.

The last test on the residuals is

$$Q^2 = n \sum_{i=1}^{k} \rho_A(i)^2 \qquad (5.12)$$

which is appr. $\chi^2(k-p)$-distributed, k is the length of the test and p is the number of estimated AR and MA terms. The test is

performed according to Li and McLeod (1981). The contribution to the χ^2-test is for lags 1, 2, 3 and 4: 2.46, 0.35, 9.64 and 34.66, respectively, giving a total of 47.11. This should be compared to a χ^2 (8,700) distribution (k•k•p=8).

5.3.5. Comparison with other models

The last part of the analysis is to estimate other ARMA-models and see if they fit data as well, and see if fewer (i.e. only one) parameters can be used.

AR(1):

$$\underline{Z}_t = \begin{bmatrix} 0.62 & 0 \\ 0 & 0.42 \end{bmatrix} \underline{Z}_{t-1} + \underline{A}_t \qquad (5.13)$$

This model uses only one parameter and the variance reduction is smaller than for the (2,0)-model. Also a larger χ^2-value(59) indicates that it is too simple a model for the actual data. AR(1)-models are often used by meteorologists to describe wind velocities.

ARMA(1,1):

$$\underline{Z}_t = \begin{bmatrix} 0.67 & 0 \\ 0 & 0.62 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} -0.08 & 0 \\ 0 & -0.25 \end{bmatrix} \underline{A}_{t-1} + \underline{A}_t \qquad (5.14)$$

The fit is not very much better than for the AR(1)-process. Furthermore, an interesting property can be observed, namely that the parameters neutralize each other so to speak. Subtracting the θ's from the ϕ's yields a model similar to the previous one. This phenomena is often observed when too many parameters are specified.

ARMA(2,1):

$$\underline{Z}_t = \begin{bmatrix} 1.04 & 0 \\ 0 & 0.37 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} -0.23 & 0 \\ 0 & 0.11 \end{bmatrix} \underline{Z}_{t-2} + \begin{bmatrix} -0.45 & 0 \\ 0 & 0 \end{bmatrix} \underline{A}_{t+1} + \underline{A}_t \qquad (5.15)$$

The neutralizing phenomenon is seen again for the north component. ϕ_{11} is much greater than expected, but in return ϕ_{12} is negative. For the east component the program eliminated θ_{21}.

Analysis showed that for the (3,0)-model the F-values of the last parameters were so low that the parameters were insignificant. This is an automatic indication of the uselessness of the last parameter.

AR(4):

$$\underline{Z}_t = \begin{bmatrix} 0.60 & 0 \\ 0 & 0.38 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} 0 & 0 \\ 0 & 0.14 \end{bmatrix} \underline{Z}_{t-2} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \underline{Z}_{t-3} + \begin{bmatrix} 0.05 & 0 \\ 0 & 0.10 \end{bmatrix} \underline{Z}_{t-4} + \underline{A}_t \qquad (5.16)$$

The fit is not better than for AR(2), what we observe is merely a more complicated model, lacking a complicated model's usual ability to reflect physically meaningful processes.

Two other models have been estimated to test the reliability of using independent estimation for two series. Dependence between the north and east components is allowed for in the following model:

$$\underline{Z}_t = \begin{bmatrix} 0.59 & -0.04 \\ 0 & 0.37 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} 0.05 & 0.02 \\ 0 & 0.12 \end{bmatrix} \underline{Z}_{t-2} + \underline{A}_t \qquad (5.17)$$

The off-diagonal elements are very small (with a corresponding low F-value) and can be excluded.

Will the model fit better if data from ST2 are included? A 4-dimensional AR(2)-model was estimated:

$$\underline{Z}_t = \begin{vmatrix} 0.59 & -0.03 & -0.01 & 0.03 \\ -0.02 & 0.36 & 0 & 0.03 \\ 0.06 & -0.34 & 0.67 & 0.04 \\ -0.04 & -0.12 & -0.02 & 0.32 \end{vmatrix} \underline{Z}_{t-1}$$

$$+ \begin{vmatrix} 0.05 & 0.02 & 0.01 & -0.01 \\ -0.01 & 0.12 & 0.04 & 0.07 \\ -0.13 & 0.18 & -0.04 & 0.03 \\ 0.03 & -0.07 & 0.09 & 0.02 \end{vmatrix} \underline{Z}_{t-2} + \underline{A}_t \ , \qquad (5.18)$$

where the four elements in the Z-vector are ST1-N, ST1-E, ST2-N and ST2-E. This approach does not change the model significantly. If $\underline{Z}_t = [U_t^1, U_t^2, V_t^1, V_t^2]^T$, and $\underline{A}_t = [a_{1,t}, a_{2,t}, a_{3,t}, a_{4,t}]^T$ then to give an example

$$U_t^1 = 0.59\, U_{t-1}^1 - 0.03U_{t-1}^2 - 0.01V_{t-1}^1 + 0.03V_{t-1}^2$$

$$+ 0.05U_{t-2}^1 + 0.02U_{t-2}^2 + 0.01V_{t-2}^1 - 0.01V_{t-2}^2 + a_{1,t} \qquad (5.19)$$

The structure of the ϕ-matrices shows that no correlation exists between ST1 and ST2, since the lower left and upper right values in the matrices are small (with one exception for ϕ_{32}, but the corresponding F-value is low). Furthermore, it shows that the appropriate coordinate system has been chosen for the given current field, since no correlation exists between the north and east components.

A rotation of the coordinate system would change the estimated model accordingly: if $\underline{\underline{C}}$ represents a transformation of $\underline{Z}$, $\underline{Y} = \underline{\underline{C}}\,\underline{Z}$, then the new $\underline{\underline{\phi}}$ matrix, $\underline{\underline{\phi}}'$, is found $\underline{\underline{\phi}}' = \underline{\underline{C}}\,\underline{\underline{\phi}}\,\underline{\underline{C}}^{-1}$.

5.4. Estimation for other series

After this analysis of ST1, week 1, m = 20 some models are shown for other data. For the sake of convenience, only the estimation part of the analysis is quoted here. First of all, to test for homogeneity and stationarity, AR(2)-models were estimated for the other stations and also for ST1 week 2. The results are shown below:

$$\text{ST2:}\quad \underline{Z}_t = \begin{bmatrix} 0.61 & 0 \\ 0 & 0.35 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \underline{Z}_{t-2} + \underline{A}_t \qquad (5.20a)$$

$$\text{ST3:}\quad \underline{Z}_t = \begin{bmatrix} 0.62 & 0 \\ 0 & 0.68 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} 0.14 & 0 \\ 0 & 0 \end{bmatrix} \underline{Z}_{t-2} + \underline{A}_t \qquad (5.20b)$$

$$\text{ST4:}\quad \underline{Z}_t = \begin{bmatrix} 0.78 & 0 \\ 0 & 0.76 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} -0.15 & 0 \\ 0 & -0.08 \end{bmatrix} \underline{Z}_{t-2} + \underline{A}_t \qquad (5.20c)$$

$$\text{ST5:}\quad \underline{Z}_t = \begin{bmatrix} 0.62 & 0 \\ 0 & 0.46 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} 0 & 0 \\ 0 & 0.07 \end{bmatrix} \underline{Z}_{t-2} + \underline{A}_t \qquad (5.20d)$$

For some of the series AR(1)-models are quite as good as AR(2)-models. The north and east models are different, but there are only minor differences between the stations, which indicates that the demand for homogeneity is satisfied.

Other models also have been estimated, but in all cases the AR(2)-models were found to provide the best fit, and are therefore the only ones shown here.

For brevity, we have chosen only the model for ST1 to represent different models for week 2:

$$\text{ST1, week 2: } \underline{z}_t = \begin{bmatrix} 0.95 & 0 \\ 0 & 0.78 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} -0.09 & 0 \\ 0 & 0 \end{bmatrix} \underline{z}_{t-2} + \underline{A}_t \tag{5.21}$$

The parameter values are higher than for week 1, but still we get a good fit using only auto-regressive terms.

The influence of the choice of m has been seen earlier, both for the spectra and for the cross-correlations. This influence is also seen in the ARMA-models, as shown below. To simplify the comparison with models we have already discussed, only the AR(2)-models are shown. First, we tried to estimate a model based on the orginal data: (ST1, week 1)

$$\underline{z}_t = \begin{bmatrix} 0.87 & 0 \\ 0 & 0.63 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} 0.12 & 0 \\ 0 & 0.30 \end{bmatrix} \underline{z}_{t-2} + A_t \tag{5.22}$$

The model does not fit too well, but that is not surprising, since it contains most of its energy at low frequencies. This is indicated by the high value of ϕ_1. It also shows us that estimating a model on the lowpass data after filtration will fail, because it will not be a stationary process.

Using highpass data, m = 100 was the first example to be analysed.

$$\underline{z}_t = \begin{bmatrix} 0.72 & 0 \\ 0 & 0.58 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} 0.21 & 0 \\ 0 & 0.28 \end{bmatrix} \underline{z}_{t-2} + \underline{A}_t \tag{5.23}$$

Already, the fit is much better than for the original series, thus proving the usefulness of filtration.

If m is changed to 50, the fit is even better.

$$\underline{z}_t = \begin{bmatrix} 0.73 & 0 \\ 0 & 0.54 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} 0.12 & 0 \\ 0 & 0.25 \end{bmatrix} \underline{z}_{t-2} + \underline{A}_t \tag{5.24}$$

Still ϕ_{11} is much higher than for the reference model m = 20. m = 10 gives a higher frequency signal than m = 20. This can be seen in the low ϕ-values.

$$\underline{z}_t = \begin{bmatrix} 0.39 & 0 \\ 0 & 0.21 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} -0.07 & 0 \\ 0 & 0 \end{bmatrix} \underline{z}_{t-2} + \underline{A}_t \tag{5.25}$$

For the east component it leads to an AR(1)-model.

Our next example will be m = 5. What remains after filtration is almost white noise, which we see in the lack of fit and the small values for the parameters. A_t will dominate the process.

$$\underline{z}_t = \begin{bmatrix} 0.09 & 0 \\ 0 & -0.04 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} -0.24 & 0 \\ 0 & -0.15 \end{bmatrix} z_{t-2} + A_t \tag{5.26}$$

The remaining step is to look at shorter periods. For week 1, ST1 a relatively calm period was found (the one also used in the correlation analysis). We estimated an AR(2)-model for this period

$$\underline{z}_t = \begin{bmatrix} 0.72 & 0 \\ 0 & 0.67 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} 0.23 & 0 \\ 0 & 0.26 \end{bmatrix} \underline{z}_{t-2} + \underline{A}_t \tag{5.27}$$

This is based on 144 observations and the fit is good ($\chi^2(8) = 24$). Although frequencies below $(24h)^{-1}$ are not present, the low-pass character of the model is obvious. The crosscorrelation

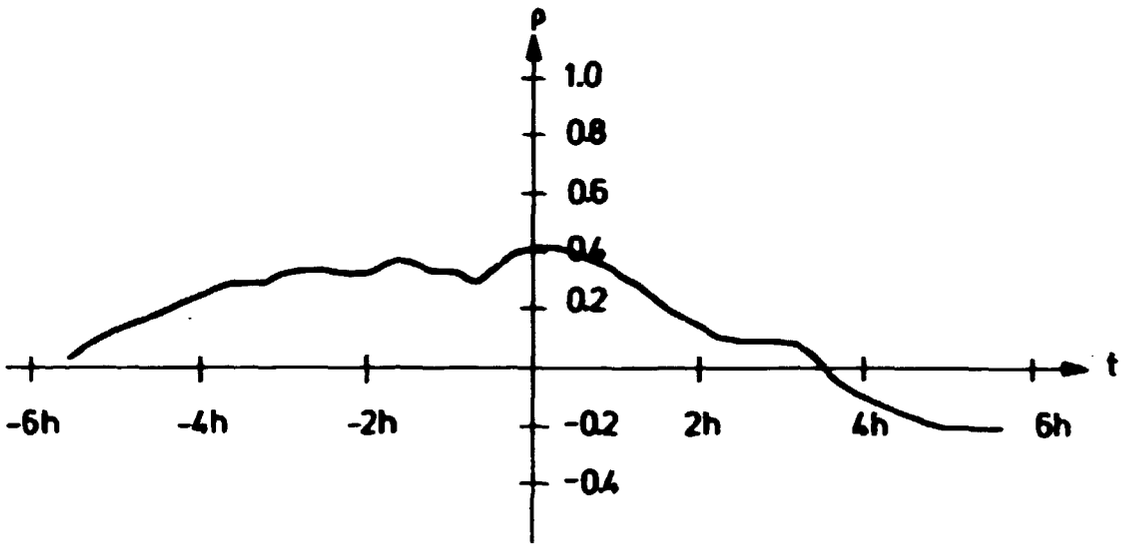


Fig. 5.9. Crosscorrelation between north and east component for St1, week 1, 144 observations.

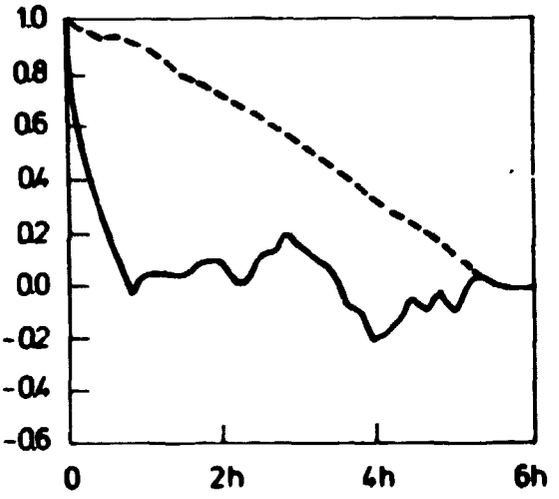


Fig. 5.10. Autocorrelations for ST1, week 1, 144 observations. ----- north component. —— east component.

function for north and east components is shown in Fig. 5.9 which could be compared with Fig. 5.2. There is still some correlation present between the two components. The ACF for the same period is seen in Fig. 5.10, where the influence from the main current is obvious in the north component. Using the highpass filter and the same time period yielded a new model

$$\underline{z}_t = \begin{bmatrix} 0.55 & 0 \\ 0 & 0.44 \end{bmatrix} z_{t-1} + \begin{bmatrix} 0.11 & 0 \\ 0 & 0.11 \end{bmatrix} \underline{z}_{t-2} + \underline{A}_t \qquad (5.28)$$

not very different from the one estimated for the whole series.

If we use other short periods we find that AR(2)-models continue to fit well, while the parameters differ significantly for the north component, the east component seems to be less dependent on the chosen time period.

In the analysis of current data from Helsingør-Helsingborg quoted in Chapter 4 a colouring of the data was made after the highpass filtration, i.e. if the variance from the lowpass part in the signal is smaller than the variance of the residuals it is necessary to add some lowpass-filtered white noise to the signal. This approach was unnecessary for our analysis of present data.

5.5. Adaptive estimation

5.5.1. Introduction

So far the discussion has taught us that an AR(2)-model will be appropriate in describing the current velocity, but we have also seen that the parameters in such a model are altered dependent on the length of the series. This indicates that for some applications a static model will be valuable, whereas it is too rigid a structure to explain all details in the current. A natural expansion is therefore to allow the parameters to change with time. The problem of estimating the parameters in such a dynamic model has many names, such as adaptive estimation, recursive

identification, adaptive filtering etc., and a variety of procedures are developed for this purpose, of which a few will be discussed briefly below.

The estimation problem can be defined as follows:

Given a series of data

$$z^t = \{z(t), z(t-1) \ldots, z(1)\} \quad , \tag{5.29}$$

we wish to infer a model of a system from the record z^t. When the model is parametrized in terms of a parameter vector θ, we say that estimation is the determination of a mapping from the data z^t to the model parameters θ:

$$z^t \rightarrow \hat{\theta}(t,z^t) \tag{5.30}$$

In adaptive estimation $\hat{\theta}(t)$ may be a function of previous data and it is therefore necessary to "condense" the observed data into an auxiliary memory quantity S(t) of fixed dimensions. This matrix will be updated after every time step, and it is this fast updating procedure that makes adaptive estimation possible. Priestly (1981) writes about recursive procedures that they "enjoy the very considerable advantage of computational simplicity." The following general scheme is used:

$$\hat{\theta}(t) = F(\hat{\theta}(t-1), S(t), Z(t)) \tag{5.31a}$$

$$S(t) = H(S(t-1), \hat{\theta}(t-1), Z(t)) \quad , \tag{5.31b}$$

where F and H are given functions.

The problem now is twofold: first, choose a suitable model parametrization; then, choose the functions F and H.

In the previous sections we have used the following notation for an ARMA-model:

$$\phi(B)Y(t) = \theta(B)e(t) \tag{5.32}$$

In the following sections we will apply the general state-space form to the dynamical system:

$$X(t+1) = FX(t) + w(t) \tag{5.33a}$$

$$Y(t) = HX(t) + e(t) \tag{5.33b}$$

where X(t) is the - in general unknown - state of the system, y(t) is the measurable output from the system, w(t) and e(t) are independent random vectors with certain covariance matrices, and F and H are matrices that usually may depend on time.

5.5.2. Application of the Kalman filter

As an example, consider an AR(2)-model in one dimension:

$$y(t) = a_1Y(t-1) + a_2Y(t-2) + e(t) \tag{5.34}$$

and use the previous notation to rewrite Eqs. (2.76) and (2.77)

$$Y(t) = F(t)\theta(t) + e(t) \tag{5.35}$$

This is called the observation equation, where Y is the set of observations. F represents the structure of the model, and θ the unobservable quantity called the state of the system, which we interpret here to be the parameter values in the AR(2)-model, i.e. $\theta = (a_1,a_2)$. F is then defined as F = (Y(t-1), Y(t-2)), and thus depends on time. The system equation, sometimes called the transition equation, reads

$$\theta(t) = F(t)\theta(t-1) + w(t) \tag{5.36}$$

where G is a known quantity; in this case G = I, the unity matrix, and w(t) is the system equation error, w(t) $\sim$ N(0,W) with W known. From this a simple Markov process arises

$$\theta(t) = \theta(t-1) + w(t) \tag{5.37}$$

One possible way to estimate the time-dependent parameters $\hat{\theta}(t)$ is to use a recursive least-square method including a loss function. The following algorithm illustrates how it is done:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[Y(t) - F(t)\hat{\theta}(t-1)] \tag{5.38a}$$

$$L(t) = \alpha_t P(t) F^T(t) \tag{5.38b}$$

$$P(t) = (1/\lambda(t))[P(t-1) - (P(t-1)F^T(t)P(t-1)^* (\lambda(t)/\alpha_t + F(t)P(t-1)F^T(t))^{-1})] \tag{5.38c}$$

where $\lambda(t)$ is the loss function (for $\lambda(t) = \lambda$ we refer to λ as the forgetting factor), and α is a weight factor usually equal to 1. The algorithm makes a least-squares estimate of the parameters, where earlier values of Y(t) become less and less important according to the loss function $\lambda(t)$. For $\lambda(t)$ constant (< 1) values of Y(t) are forgotten after $1/(1-\lambda)$ time step.

Using $\lambda = 1$ gives an estimation for θ for the whole series, which could be compared to estimates from the MARIMA programme. But the advantage of the algorithm is that by allowing the parameters to change, a more reliable picture of the current can be given.

One of the most useful tools in making an adaptive estimation is the Kalman filter, originally described by Kalman (1960) and Kalman and Bucy (1961) and now quoted in many papers and textbooks (Ljung & Søderstrøm (1983), Harvey (1984)). In Sec. 2.2.6 we introduced the topic, here a practical example will be given.

The determination of the best estimate of $\theta(t+1)$ given observations Y^t up to time t, which is the conditional mean $\hat{\theta}(t+1) = E(\theta(t+1)|Y^t)$, is performed in the following way (Ljung and Søderstrøm (1983)).

$$\theta(t+1) = F(t)\hat{\theta}(t) + K(t)[Y(t) - G(t)\hat{\theta}(t)] \tag{5.39a}$$

$$K(t) = F(t)P(t)G^T(t)[R_2(t) + G(t)P(t)G^T(t)]^{-1} \tag{5.39b}$$

$$P(t+1) = F(t)P(t)F^T(t) + R_1(t) - K(t)G(t)P(t)F^T(t) \tag{5.39c}$$

The dimensions of the matrices are as follows:

Y: (mx1), θ: (nx1), F: (mxn), G: (nxn)

R_1: (m×m), R_2: (n×n), K: (m×n), P: (n×n)

where m is the number of simultaneous observations, n the number of parameters, and R_1 and R_2 are the covariance matrices for w and e, respectively.

For a one-dimensional AR(2)-model the procedure is straightforward, and the results are shown in the next section. For a two-dimensional AR(2)-model

$$\underline{Y}(t) = \begin{vmatrix} \phi_{11,1} & \phi_{12,1} \\ \phi_{21,1} & \phi_{22,1} \end{vmatrix} \underline{Y}(t-1) + \begin{vmatrix} \phi_{11,2} & \phi_{12,2} \\ \phi_{21,2} & \phi_{22,2} \end{vmatrix} \underline{Y}(t-2) + \underline{e}(t) \tag{5.40}$$

we get the following state-space form (using for convenience θ instead of ϕ as the AR-coefficients)

$$\begin{bmatrix} \theta_{11,1}(t+1) \\ \theta_{12,1}(t+1) \\ \theta_{21,1}(t+1) \\ \theta_{22,1}(t+1) \\ \theta_{11,2}(t+1) \\ \theta_{12,2}(t+1) \\ \theta_{21,2}(t+1) \\ \theta_{22,2}(t+1) \end{bmatrix} = \begin{bmatrix} \theta_{11,1}(t) \\ \theta_{12,1}(t) \\ \theta_{21,1}(t) \\ \theta_{22,1}(t) \\ \theta_{11,2}(t) \\ \theta_{12,2}(t) \\ \theta_{21,2}(t) \\ \theta_{22,2}(t) \end{bmatrix} + \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ w_4(t) \\ w_5(t) \\ w_6(t) \\ w_7(t) \\ w_8(t) \end{bmatrix} \tag{5.41a}$$

$$\begin{vmatrix} Y_1(t) \\ Y_2(t) \end{vmatrix} =$$

$$\begin{vmatrix} Y_1(t-1) & Y_2(t-1) & 0 & 0 & Y_1(t-1) & Y_2(t-2) & 0 & 0 \\ 0 & 0 & Y_1(t-1) & Y_2(t-1) & 0 & 0 & Y_1(t-1) & Y_2(t-2) \end{vmatrix}$$

$$\times \begin{vmatrix} \theta_{11,1}(t) \\ \theta_{12,1}(t) \\ \theta_{21,1}(t) \\ \theta_{22,1}(t) \\ \theta_{11,2}(t) \\ \theta_{12,2}(t) \\ \theta_{21,2}(t) \\ \theta_{22,2}(t) \end{vmatrix} + \begin{vmatrix} e_1(t) \\ \\ e_2(t) \end{vmatrix} \qquad (5.41b)$$

With reliable assumptions for the covariance matrix, R_1 for w and R_2 for e, the Kalman filter can be implemented directly, giving an adaptive estimation of the parameters.

5.5.3. Results

The RLS was used in the estimation for a one-dimensional series, viz. ST1, week 1, north component after highpass filtering, and the model estimated was an AR(2)-model. The two parameters were estimated at $\phi_1 = 0.57$ and $\phi_2 = 0.05$ in agreement with results from the MARIMA-program (Sec. 5.3). Then the forgetting factor, λ, was introduced, and different values were used. For $\lambda = 0.95$ relatively slow variations are found, but the range of the parameters is high, for ϕ the lowest value found is 0.27 whereas the highest is 1.2. For ϕ_2 the value goes from -0.37 to 0.34. (The two parameters depend of course, on each other). This result is an indication of the time-changing qualities of the current velocity and shows the need for time-varying models. One reason for this need concerns the nature of the mean current, the direction of which is restricted to be either northward or southward.

Also a Kalman filter was constructed for the one-dimensional AR(2)-model, and the results from this can be compared to the aforementioned RLS algorithm. The "memory" of the Kalman filter is specified indirectly via the parameter values and their expected variance (R2 in the algorithm). ϕ_1 was in the range 0.3 - 1.02 and ϕ_2 was between -0.28 and 0.26. A direct comparison shows that the two estimation methods do not differ significantly for an AR(2)-process.

For the eight-dimensional ϕ-vector a similar Kalman filter was constructed thus giving the time variations for the eight parameters defined in Eq. 5.41. A few examples of results are shown in the figures that follow. Some aspects of the estimation will be discussed below.

The choice of R1 and R2, the covariance matrices for the noise processes, was not seen as influencing the parameter values significantly. An example for two parameters ($\phi_{11,1}$ and $\phi_{22,1}$) are seen in Fig. 5.11. In the beginning there are some estimation troubles, whereas after some time only slow variations in the parameters occur. For $\phi_{22,1}$ there is a sudden decrease at about observation 400, which may be explained by an unstable current direction. The difference between the north and east direction is seen by the larger value for ϕ_{11} indicating the importance of the mean current direction, which led to slower variations in the N/S direction.

For all stations it was seen that the off-diagonal parameters (ϕ_{12} and ϕ_{21}) were small and oscillating around zero, in accordance with earlier result showing the reason for eliminating these parameters. Comparing $\phi_{11,1}$ and $\phi_{11,2}$ expresses the problem of redundancy. $\phi_{11,2}$ is also very low and depends on $\phi_{11,1}$. Therefore an AR(1)-model was estimated; it was seen that doing this scarcely changed the estimates. This analysis was made in two ways: A visual examination of plots of the type given in Fig. 5.11 was made, and the autocorrelations and crosscorrelations for the parameters were determined.

The values found can be compared to results from MARIMA, and good agreement was found.

Furthermore, the residuals have been analysed, plots were made, spectra were drawn, and auto- and crosscorrelations were calculated. The "whiteness" of the residuals was evident, as seen from the spectra and autocorrelation functions. The independence between residuals from the north and east series also became clear, this was true as well regarding the independence between corresponding residuals from different stations. It was also seen that the amount of variation explained by the model of course is significantly larger for a model where the parameters are allowed to change with time.

With the problems of finding a correlation structure described in Chapter 4 kept in mind, a correlation analysis was made on the parameter values. A direct comparison of some of the values can be made using Fig. 5.11 and 5.12, and it is seen that some correlation must exist between stations, at least for ϕ_{11}. But

Fig. 5.11. Estimation of $\phi_{11,1}$ (PAR1) and $\phi_{22,1}$ (PAR4) (compare Eq. 5.40) for an AR(2)-process using a Kalman filter, ST4.

checking the crosscorrelation functions for different combinations did not provide a clearer picture than was already found earlier.

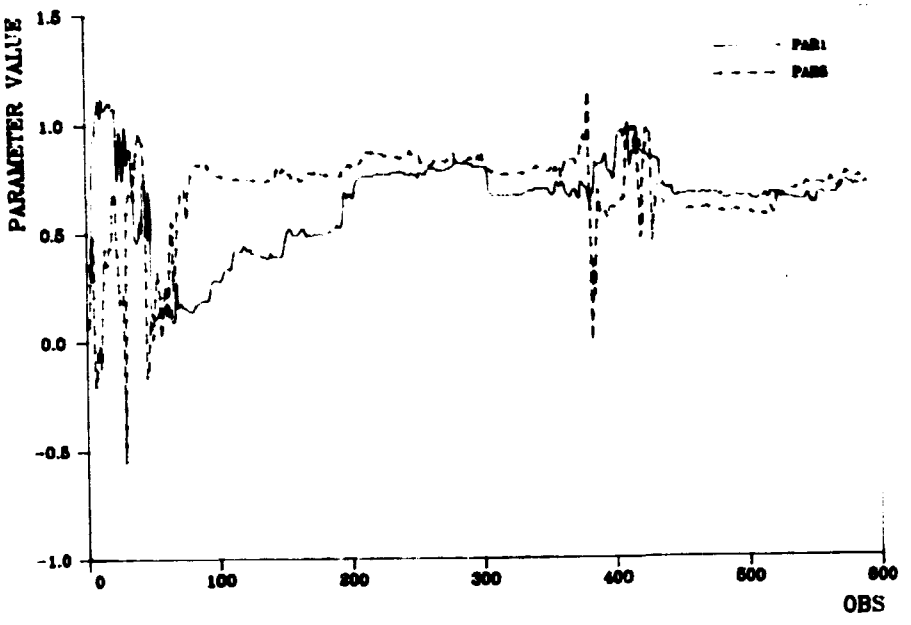


Fig. 5.12. Estimation of $\phi_{11,1}$ in an AR(2)-process for ST1(PAR1) and ST2(PAR5) using a Kalman filter.

5.5.4. Discussion

This section has shown the benefit in using models with time-varying parameters. On using a Kalman filter in estimating the parameters in autoregressive models we arrive at a much more detailed picture of the current velocity measured at the five stations. The residuals from such models are much smaller than those arising from estimation by MARIMA, and new correlation analyses can be made. Although the variance is more important than the parameter values for the simulation model (Chapter 6), time-changing parameters will give a more reliable dilution process than using constant values.

Here only AR-models have been analysed, but many other forms can be suggested (Harrison and Stevens, 1976), (Harvey, 1983). The following model represents one:

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{5.42}$$

where μ_t is a time-varying mean, γ_t can be a seasonal component and ε_t is the noise. This model can easily be estimated using the Kalman filter. The γ_t could also represent the rate of change in the mean level, and still the Kalman filter could be used. Also ARIMA-models can be handled by the Kalman filter.

A further analysis of current data should include Kalman filtering of such model structures.

5.6. Conclusion

In this chapter ARMA-models for current velocities have been estimated. For the velocity at ST1 measured during week 1 and highpass filtered with a cutoff frequency of $(7h)^{-1}$, the following AR(2)-model was seem to fit data best

$$\underline{z}_t = \begin{bmatrix} 0.59 & 0 \\ 0 & 0.37 \end{bmatrix} \underline{z}_{t-1} + \begin{bmatrix} 0.05 & 0 \\ 0 & 0.12 \end{bmatrix} \underline{z}_{t-2} + \underline{A}_t \tag{5.43}$$

This model is used in the next chapter to simulate particle movements. We have studied many other ARMA-models for the same data and also estimated models on other data, such as other stations, other time periods, and using other filtrations. Using a Kalman filter the parameters have been estimated recursively. In general, AR(2)-models were convenient to use. Many similarities were found between the models, thus confirming the homogeneity of the area. No crosscorrelation was found between the north and east series after filtration and therefore the models were estimated independently.

6. THE SIMULATION MODEL

6.1. Introduction

"The temptation for man to make guesses at the future course of events has always been strong and draws its origins from astrology, prophecy, and soothsaying" (Bennett, 1979). The forecasting method used in the present study is the Box-Jenkins analysis of time series. It is used to predict the dilution of effluents in the marine environment. The simulation model is constructed especially to illustrate the dilution of discharges of radioactivity from the Swedish nuclear power plant Barsebäck in the Sound, but we hope that the principles outlined here can be extended to other places. Diffusion of other kinds of pollution (e.g. oil spills) can also be simulated using the proposed technique. For each area under consideration, new data must be collected to describe the recipient.

The principles for a simulation model based on current measurements have been described earlier by Boelskifte (1980) and Spliid et al. (1981). An extended method of simulating simultaneous particle movements in water are outlined in this chapter. ARMA-processes are used and their mutual correlation, which depends on displacement, is one of the most important parameters. Data used in the model have now been tested in several ways (Chapters 4 and 5). The correlation structure of the area has been studied in detail. In order to estimate the longitudinal and transverse correlation functions, the effects of highpass filtering have also been studied. The assumptions for the model such as homogeneity, isotropy and stationarity have been tested. Furthermore, the model's output is now in accordance with the statistical formulation in Section 2.1. A further development of the model is seen in Section 4, where a conditioned simulation is performed, and in Section 6 of this chapter, where an adaptive model is used. Chapter 7 includes an example of how the use of bioindicators has made it possible to verify the model.

6.2. Assumptions

In order for the statistical model to represent real physical processes, the current field must satisfy some restrictive assumptions. First of all, homogeneity is required, i.e. the statistical description for the current must be equal for all points in the area under consideration. This is confirmed by the current measurements since similar ARMA-models can be fitted to the different stations and the mean and variance are also similar. Another important property is isotropy, or the situation where the statistical characteristics are equal, regardless of the direction considered. This is discussed in Section 5.4. Stationarity is the third demand to meet. In other words, moments of any order of the current velocity must be invariant with time. The estimated AR(2)-models confirm the presence of stationarity, since the roots in the characteristic polynomium $1-\phi_1 z-\phi_2 z^2 = 0$ lie outside the unit circle (for $\phi_1 = 0.59$ and $\phi_2 = 0.05$ the roots are, respectively, -13.3 and 1.5).

The last assumption used in the model is that the longitudinal correlation, f, is an exponentially decreasing function of distance, $f(r) = \exp(-r/L)$. In case of isotropy, the transverse correlation, g, can be deduced from this using equation 2.4, which gives $g(r) = (1-r/2L)\exp(-r/L)$. Although this form of f and g could not be seen clearly from the current measurements, it seems to be an intuitively reliable assumption, since some inverse proportionality must exist between distance and correlation. Thus only one parameter, the characteristic length scale L, is necessary to represent the correlation structure. From the analysis in Chapter 4 it is seen that 900 m can be an estimate of L, since only low correlations are found for this distance, when the high-pass filtered data are used. The influence of the choice of L is tested in Section 6.5.

As we wrote in Chapter 5, the following AR(2)-model will be used:

$$\underline{Z}_t = \begin{bmatrix} 0.59 & 0 \\ 0 & 0.37 \end{bmatrix} \underline{Z}_{t-1} + \begin{bmatrix} 0.05 & 0 \\ 0 & 0.12 \end{bmatrix} \underline{Z}_{t-2} + \underline{A}_t \qquad (6.1)$$

This is estimated using the high-pass filtered data where $(7h)^{-1}$ is used as cut-off frequency.

The problem of Lagrangian/Eulerian description has been discussed earlier (Section 2.1). One conclusion is that it could be reduced to a matter of time scale. It is also a fundamental problem for this study, but we assume the two frames of reference to be equal, i.e. Eulerian measurements are used to simulate Lagrangian movements. In this way, only a slight error is introduced, since the degree of turbulence is small.

6.3. Description of the model

In simulating the movement of a single particle, we use the two independent AR(2)-models for north and east velocity. Thus a picture like Fig. 6.1 can be drawn, where we illustrate the principle for simulating an ARMA-process (compare Section 2.2).

Adding a new particle creates a band on the random shocks (the A_t's) which now should be correlated to an extent dependent on the distance between the particles, because the correlation between the Z_t's for two realisations of the same ARMA-model equals

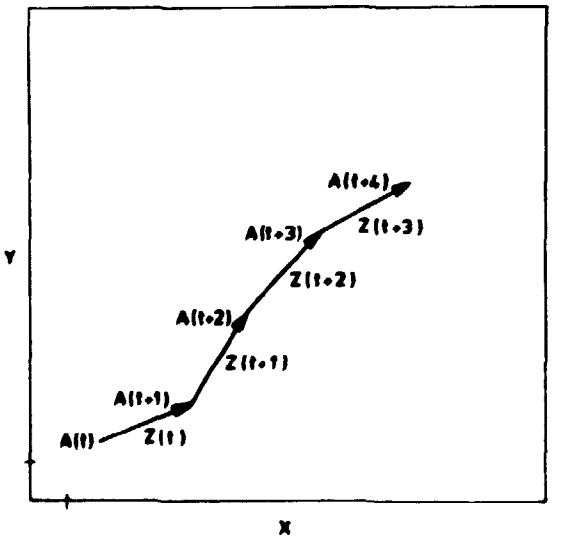


Fig. 6.1. Simulation of movements of a single particle.

the correlation between their random shocks. Therefore, in the simulation model the correlation between velocities is obtained by introducing a correlation in the A_t's, which are used as input to the processes.

To express the correlation between two particles it is necessary to refer to a fixed coordinate system, in this case the north east system. Consider two points, P and P', a distance r apart and noise vectors A and A' respectively, as shown in Fig. 6.2. Our earlier assumptions imply that the correlation between A and A' can be expressed in terms of $cor(A_{long},A'_{long}) = f(r)$ and $cor(A_{norm},A'_{norm}) = g(r)$, the longitudinal and transverse correlation, respectively. It is assumed that $var(A_{long}) = var(A_{norm})$ and that $cor(A_{norm},A_{long}) = cor(A'_{norm},A'_{long}) = 0$.

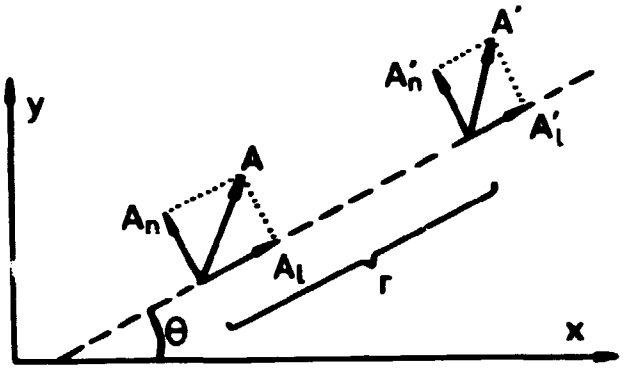


Fig. 6.2. Generation of spatially correlated white noise. (From Spliid et al. (1981).)

Relative to a fixed x-y coordinate system the correlation between A and A' is given in the following correlation matrix, $\underline{\Sigma}$:

$$\text{cor}\begin{vmatrix} A_x \\ A_y \\ A'_x \\ A'_y \end{vmatrix} = \begin{vmatrix} 1 & 0 & fc^2+gs^2 & (f-g)sc \\ 0 & 1 & (f-g)sc & fs^2+gc^2 \\ fc^2+gs^2 & (f-g)sc & 1 & 0 \\ (f-g)sc & fs^2+gc^2 & 0 & 1 \end{vmatrix} \qquad (6.2)$$

where $f = f(r)$, $g = g(r)$, $c = \cos(\theta)$, and $s = \sin(\theta)$.

Knowing the correlation which the noise processes are supposed to satisfy, the following procedure is used to find the A's: One finds the eigenvectors $\underline{P}_k$ and eigenvalues λ_k for Σ. Then the noise vector, $\underline{A}$, to be used as input to the ARMA-model is found as

$$\underline{A} = (\underline{P}\ \underline{\Lambda}^{1/2})\underline{U} \qquad (6.3)$$

where $(\underline{P}\ \underline{\Lambda}^{1/2})$ is shown to be

$$\underline{P}_1\ \lambda_1^{1/2}\ ,\ \underline{P}_2\ \lambda_2^{1/2}\ ,\ \ldots\ldots\ ,\ \underline{P}_n\ \lambda_n^{1/2}$$

and $\underline{U}$ $N(\underline{O},I)$ is a vector which can be created by a random number generator.

Using the AR(2)-model a simulation of particle movements can be made. For two particles an example is seen in Fig. 6.3. This procedure explains why the simulation model is restricted to relatively few particles. This is so because finding eigenvectors for a matrix involves producing its inversion, a process that is very timeconsuming if the matrix is large. But there are no theoretical problems presented by the use of more than two particles. The dispersion matrix is just extended accordingly.

Using three particles allows for the output from the model to be illustrated if ellipses are drawn between the three points. For each simulation a picture of the size of the cloud can be drawn. An example of this is seen in Fig. 6.4, where the ellipses drawn are the smallest ones that can be constructed covering the three points. The time interval used in the model is 10 minutes as it is for the current measurements. For each 18 time steps (3 hrs) ellipses are plotted.

Looking at three points and one ellipse gives rise to different ways of expressing the dispersion. The distance between points can be used as a figure for relative dispersion. The change in position of the centre of gravity of the three points can be

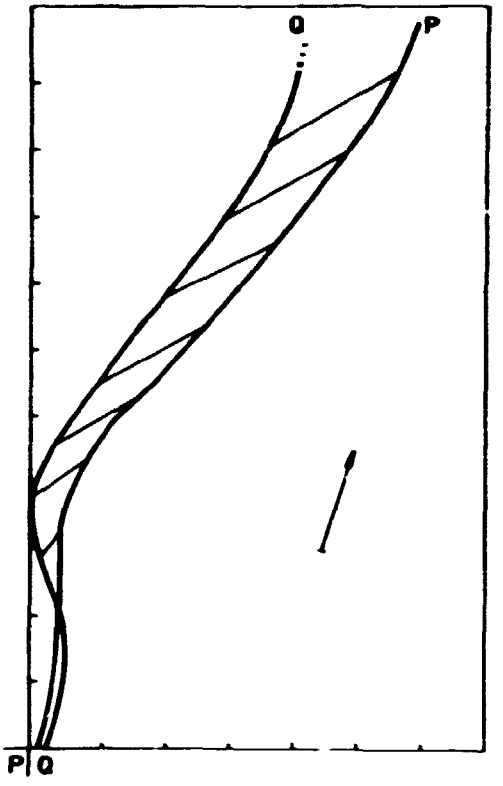


Fig. 6.3. Simulation of the simultaneous movements of two particles.

studied, and the area of the ellipse can also be used in expressing the dynamics of the turbulence. But the last figure has one major drawback. If the three points are located along a straight line, the ellipse will be very narrow, and consequently the area will be small. The area is expressed as πab, where a and b are the semi-axes of the ellipse. Instead, we use an "area-equivalence" figure to illustrate the area expressed as the sum of the squared semi-axes. These can be found as the eigenvalues for the matrix D(X,Y) expressing the moment of inertia of the triangle. It is found as

$$D(X,Y) = \begin{vmatrix} X1 & X2 & X3 \\ Y1 & Y2 & Y3 \end{vmatrix} \begin{vmatrix} X1 & X2 & X3 \\ Y1 & Y2 & Y3 \end{vmatrix}^{T} \qquad (6.4)$$

where X1, X2, X3, Y1, Y2, Y3 are coordinates in the coordinate system using the gravity as origon.

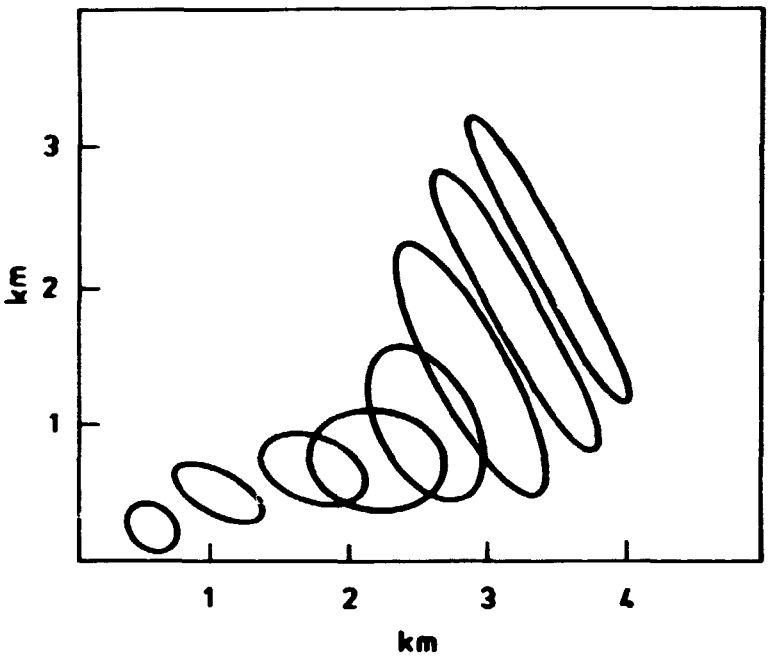


Fig. 6.4. Simulation example with the three-moving-points model. The time interval between ellipses is three hours.

Merely by taking the inverse one can use this "area-equivalence" to express the dilution directly. Another advantage of this area-equivalence is that it is not so sensitive to those situations that often occur in simulations where two of the three points are relatively close together. Therefore, it is used intensively in the following.

6.4. Results

Our aim is to find the average of the aforementioned area-equivalence for many simulations and covering a time span as long as as properly. This is illustrated in Fig. 6.5, where the average of 100 simulations as well as two examples of single simulations are drawn. The average value is an indication of the most probable dilution found after a given time period following the discharge. It expresses the same qualities as $E\{\sigma_x^2\}$ in equation (2.13) - the ensemble average of the variance of each cloud about its center of mass.

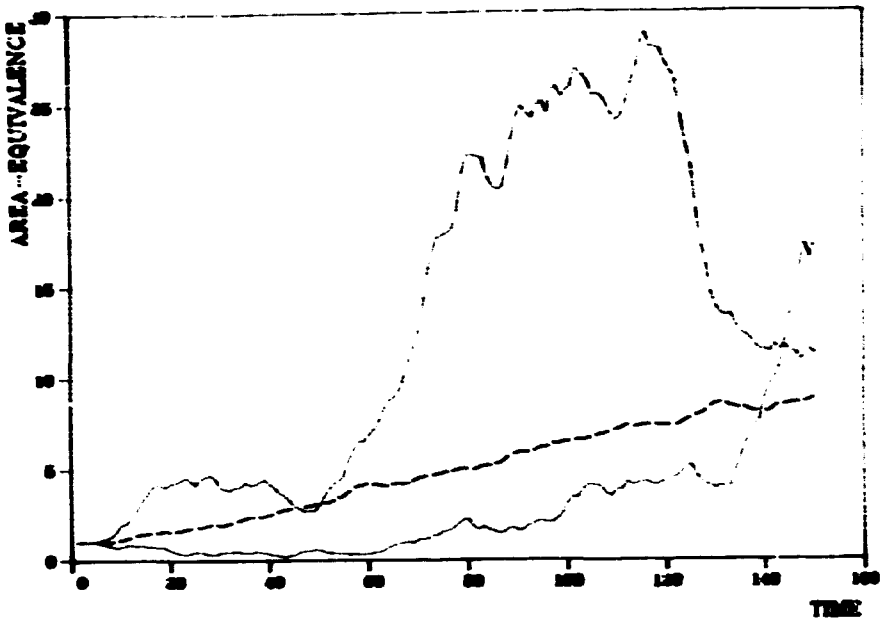


Fig. 6.5. a) Average of 100 simulations (----)
and two examples of independent simulations (——).
144 timesteps represent 24 hours.

For each simulation we find how the center of mass of an individual cloud is displaced from its expected position. From this the ensemble mean-square displacement can be calculated, as can the variance of the ensemble distribution about its expected position σ_x^2. The two single simulations show how different the simulations are and thereby stress the importance of using average values.

From the estimated AR(2)-model and the correlation structure, we can determine the dilution process. In deciding how the simulation is to be performed, we must determine what is interesting to study and what is most appropriate to running the simulations. Therefore, if we use three particles in the simulations we do not gather more information about the area than if we used only two. However, it does make it much easier to illustrate the process by an example.

It is of interest to know not only the average value of many simulations but also the variance of this value. It serves as a measure of the influence of the choice of the white noise processes. If this variance is high, the model structure plays a minor role compared to the white noise. It is not an indication of the degree of turbulence, which is expressed indirectly in the area-equivalence; rather, it indicates the amount of information contained in the model compared to the stochastic element. The minimum and maximum values for each time step have the same property; they merely indicate the differences between two simulations more directly. These values we find for each timestep, which means that they are not necessarily taken from the same simulation for each timestep.

Tabel 6.1. Area-equivalence and distance from (0,0) to center of mass together with their minimum and maximum values for the standard example including 100 simulations.

Time hours	Relative area-equivalence			Distance to (0,0) meters		
	mean	min	max	mean	min	max
0	1.0	1.00	1	0	0	0
2	1.6	0.08	3.8	188	5	489
4	2.5	0.04	8.9	303	18	584
6	3.2	0.11	11.9	382	76	858
8	4.2	0.09	18.6	396	16	983
10	4.8	0.17	19.8	457	25	1260
12	5.7	0.03	23	515	37	1500
14	6.9	0.02	28	541	51	1520
16	7.7	0.09	42	598	131	1540
18	9.4	0.06	40	648	28	1590
20	11.3	0.08	52	674	105	1930
22	12.7	0.04	81	706	32	1950
24	13.9	0.04	105	749	98	2140

Some of the results are shown in Table 6.1. For different time-steps the mean and maximum and minimum values are shown for the area-equivalence and the distance to the point (0,0).

The minimum and maximum values show how different the simulations can be. With a single simulation nothing is known about the corresponding probability, whereas the average value should represent the most frequently found situation. One may ask, therefore, how many simulations are necessary to give a reliable model of the delution process. A single example of 365 simulations has shown us that the differences between results from 100 and from 365 simulations are not significant.

The very small minimum values for the area-equivalence arises when two particles come close together. This statistical phenomemon will be seen sometimes out of 100 simulations. For the average value it is not significant whether the minimum value or the number one is used.

In Chapter 3 we saw that in the initial stage the size of the cluster would be proportional to T, whereas in the inertial subrange it would be proportional to $T^{2/3}$. Results from the simulation experiment can be interpreted in the same way. With the chosen initial conditions we get a proportionality close to T. But this result is very sensitive to the choice of initial separation between particles and the variance for the white noise processes. For some combinations of initial conditions we have found proportionality to $T^{1.3}$. This is more in accordance with our expectations, since the model assumes that the processes take place in the inertial subrange.

6.4.1. Conditional simulation

Until now the analysis has been based on a statistical description alone. It is possible to expand the model so that instead of using randomly generated data as input to all 6 series, the actual measured high-pass filtered velocity is used for one of the three particles. Knowing the velocity and the AR(2)-model, the random shocks can be found ($A_t = Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2}$). For

three points (6 velocities) 2 velocities are given, but all 6 have to satisfy the dispersion matrix given according to the correlation. This means that the usual way of finding the A_t's cannot be used, since the two known A_t's make a bond on the other four.

From Eq. 6.3 it can be seen that A_1, for example, is found as

$$A_1 = U_1P_{11} + U_2P_{12} + U_3P_{13} + U_4P_{14} + U_5P_{15} + U_6P_{16} \qquad (6.5)$$

where P_{ij} is element i,j in the matrix $(\underline{\underline{P}}\ \underline{\underline{\Lambda}}^{1/2})$. For instance, if the actual values from ST1 are used for A_5 and A_6, then the other four processes A_1, A_2, A_3, and A_4 will follow a conditional probability function. Let the vector of noise input be

$$\underline{X} = \begin{matrix} \underline{X}_1 \\ \underline{X}_2 \end{matrix} \quad , \quad \text{with dispersion matrix} \qquad (6.6a)$$

$$\underline{\underline{\Sigma}} = \begin{matrix} \underline{\underline{\Sigma}}_{11} & \underline{\underline{\Sigma}}_{12} \\ \underline{\underline{\Sigma}}_{21} & \underline{\underline{\Sigma}}_{22} \end{matrix} \qquad (6.6b)$$

where $\underline{X}_1 = (A_1, A_2, A_3, A_4)^T$

and

$$\underline{X}_2 = (A_5, A_6)^T$$

Then the conditional mean value of $\underline{X}_1$ is given by - assuming that the unconditional mean for all elements is zero -

$$E(\underline{X}_1 | \underline{X}_2 = (A_5, A_6)^T) = \underline{\underline{\Sigma}}_{12}\ \underline{\underline{\Sigma}}_{22}^{-1}(A_5, A_6)^T \qquad (6.7)$$

and the conditional dispersion matrix of $\underline{X}_1$ is given by

$$D(\underline{X}_1 | \underline{X}_2 = (A_5, A_6)^T) = \underline{\underline{\Sigma}}_{11} - \underline{\underline{\Sigma}}_{12}\ \underline{\underline{\Sigma}}_{22}^{-1}\ \underline{\underline{\Sigma}}_{21} \qquad (6.8)$$

For each timestep $\underline{\underline{\Sigma}}$ is found, and thereby $D(X_1)$ can be determined. Using (A_5, A_6) the addition to A_1, A_2, A_3, and A_4 can easily be calculated.

This procedure is for the high-pass part; to follow the real current the low-pass part should be added to all velocities. Now a picture of a real situation can be given.

Examples are seen in Table 6.2. Also here some of the minimum values are low. The maximum values and the area equivalence are much lower than for the unconditional examples. This is due to the use of the same track for one of the particles of every simulation. It makes very long distances between particles less probable.

Tabel 6.2. Area-equivalence and distance from (0,0) to center of mass together with their minimum and maximum values for conditional simulation.

Time hours	Relative area-equivalence			Distance to (0,0) meters		
	mean	min	max	mean	min	max
0	1	1	1	0	0	0
2	1.2	0.8	2.2	43	6	20
4	1.2	0.6	2.5	60	12	147
6	1.2	0.6	3.2	59	5	153
8	1.6	0.7	4.4	96	21	172
10	1.2	0.5	4.6	77	13	173
12	1.2	0.3	4.0	87	9	270
14	1.3	0.4	4.5	96	10	278
16	1.3	0.4	4.8	100	12	288
18	1.3	0.4	5.1	107	6	280
20	1.4	0.3	5.8	114	5	300
22	1.5	0.2	5.7	130	21	332
24	1.6	0.4	6.2	129	28	385

Nevertheless, the qualitative properties of the two methods are equal. The advantage of this last method lies in the fact that it allows us to follow an actual current situation in more detail.

6.4.2. Drogue experiments

Two drogues were constructed and used in the investigation as a supplement to the current measurements. The purpose was twofold: First, to have a realisation of a relative diffusion experiment which could be compared to results from the simulation model. Secondly, to produce Lagrangian data to compare with the Eulerian data received from the current measurements.

The drogues were constructed so the depths could be changed. The ones we used were placed at about 6 m, approximately the same depths at which the current measuring instruments were placed. The influence of the wind can also be ignored.

The drogues were used on two different days; one during week 1, and the other during week 2. Each experiment lasted for about 6 hours (i.e. 36 time steps in the simulation model). Too few data were produced by this experiment for a detailed analysis of relative turbulent diffusion, as described in Chapter 2. But nevertheless, the results can be compared with those found in the simulation model for the standard example.

To compare the two descriptions (Eulerian and Lagrangian) the track was deduced for a particle following exactly the current measured at ST1. There was no agreement between the two. The drogues were drifting towards the west; whereas the current meters showed a northerly current. We looked for an explanation in SMHI, (1977) where there is a picture showing a common current situation in the Lundåkra Bay. According to SMHI the current is westerly near an east-west line from Barsebäck. This means that the homogeneous area where the current meters were placed is limited to the south. After some time when the drogues reached this line they followed the current towards west and thereby produced data which differed from what have been expected from the current measurements.

6.5. Sensitivity analysis

This section presents some results from the simulation model with changes in different parameters. The results are compared to the aforementioned examples.

Only some aspects of the model can be compared, but hopefully they will provide an idea of the direction a change in one single parameter would result in. An analytical analysis of sensitivity will not be given.

We know the variance of the residuals (A_t's) from the estimation procedure. This variance is expected to be a crucial parameter, since the A_t's are the inputs of the process. In Table 6.3, values of the area-equivalence for different values of σ^2 are seen. There it is confirmed that the influence is highly significant. This is not surprising, since it is an indication of the degree of turbulence. Therefore, this parameter was studied more intensively in the Section about making adaptive estimates.

The other critical parameter is the characteristic length scale ℓ, and therefore the correlation functions f and g. No analysis is made here concerning the form of f and g; only the value of ℓ is investigated. Examples are found in Table 6.3, where it is seen how important this ℓ value is. On the average, ℓ will change the simulated dilution process qualitatively, for when the separation between particles equals ℓ, i.e. there is no correlation between particles, the rate of diffusion increases.

This influence is confirmed by the simulations and it justifies the use of so many manipulations to find the "right" value of ℓ. (what exactly the "right" value is, is still, to a certain extent, an open question).

The boundary conditions, i.e. the initial separation between particles, is also important. The "right" choice depends on the kind of process we want to investigate. When can we say the process begins? Photographs of temperature rises in the plume from Barsebäck (SMHI, 1976) can give an idea of the initial width of the plume.

Table 6.3. Sensitivity analysis on the relative area-equivalence. See text for explanation.

Time hrs	a	b	c	d	e	f	g	h	i	j
0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2	1.6	1.1	4.0	2.2	1.4	1.2	2.7	2.2	1.5	1.6
4	1.9	1.2	7.4	2.7	1.5	1.3	4.1	3.3	1.7	2.0
6	2.7	1.2	12.6	3.6	1.8	1.5	6.2	5.0	2.2	2.4
8	3.4	1.3	18.1	5.1	2.1	1.6	7.5	7.1	2.5	3.0
10	4.2	1.4	22.7	6.2	2.5	1.8	9.8	9.3	2.7	3.4
12	5.0	1.4	30.7	7.2	3.2	1.9	12.2	10.5	3.0	3.9
14	5.8	1.5	36.7	8.1	4.0	2.0	13.3	11.0	3.5	4.8
16	6.6	1.5	44.4	9.9	4.5	2.3	16.1	14.4	3.6	5.4
18	7.3	1.5	47.6	11.3	5.1	2.5	19.5	16.0	4.1	6.0
20	7.9	1.5	54.9	13.0	5.7	2.4	21.5	16.6	4.2	6.6
22	8.1	1.5	61.3	13.4	6.3	2.6	25.3	18.0	4.6	7.0
24	8.8	1.6	68.5	13.6	6.6	2.6	25.8	19.4	5.2	7.4

a) A standard example
b) $\sigma^2 = 1$
c) $\sigma^2 = 8$
d) $\ell = 500$
e) $\ell = 1800$
f) Initial distance between particles: 900 m
g) Initial distance between particles: 150 m
h) North: $\phi_1 = 0.7$, $\phi_2 = 0.05$, East: $\phi_1 = 0.5$, $\phi_2 = 0.12$
i) North: $\phi_1 = 0.4$, $\phi_2 = 0.05$, East: $\phi_1 = 0.2$, $\phi_2 = 0.12$
j) North: $\phi_1 = 0.59$, $\phi_2 = 0.0$, East: $\phi_1 = 0.37$, $\phi_2 = 0.0$

200 - 300 m between particles seems to be representative. The importance of this separation depends on whether the absolute level or the rate of dilution is of interest.

The estimated AR-model is perhaps the most interesting part to investigate, since the ϕ-parameters represent a high degree of certainty as compared to, for instance, the correlation function. Many examples could be given, but here we restrict the analysis to three situations: 1) exlude the second AR-term, 2) increase both variables, and 3) decrease them. The results of these experiments are also shown in Table 6.3.

The importance of the parameter values is significant when compared to the influence of the variance of the process, and the characteristic length scale. The influence of larger changes in the parameters is highly significant.

As stated earlier a single example has shown that using 365 simulations instead of 100 does not change the results significantly.

6.6. Discussion

After having tested the statistical properties of the model it is important to compare results with physical measurements and turbulence theories. The drogue experiment is an example of a feasible test. However, it did not fulfil our expectations, due to inhomogeneities in the nearby environment. The relative diffusion, i.e. the development in distance between the two drogues, was not compared to the models output, since the trajectories did not represent the current measurements quite well.

The other physical test which we performed was to collect seaweed and use it as an indicator of the dilution of ^{60}Co released from the Swedish nuclear power plant Barsebäck. This test is described in detail in Chapter 7, where it is seen that the concentration, C, in seaweed is significantly described as a function of the distance, X, from Barsebäck as such: $C = aX^{-1.4}$ (for explanation and a further discussion see the next chapter).

How reliable are the data used in the model? The knowledge of the correlation function is still insufficient, although minor changes influence the model only slightly. The degree of turbulence is the

most important parameter which we know in detail from the estimation procedure, the subjective choice concerns only the length of the high-pass filter. Also, the parameters used in the AR-model represent a high degree of certainty. However, it should be remembered that all the parameters used here are valid only for the area near Barsebäck. Any new application of the model requires a new measuring programme.

Such a programme should include measuring the vertical current profile. If possible, the correlation structure should be studied more intensively. In order to do this, more current meters should be used in the investigation area. A significant way to improve the model would be to use three dimensions, since the shear is a significant factor in the dispersion process. Longer series could be used to extend the adaptive estimati[illegible] thereby improving the reliability of the simulation of long-term effects. This should be done using the conditional simulation technique as described earlier.

6.7. Conclusion

This chapter contains a description of the model simulating relative diffusion. It has been possible to get results that, compared to turbulence theory and practical measurements, describe the diffusion process in a reliable manner. The assumptions related to the model are outlined together with a more detailed explanation of the theoretical background applied in the simulation scheme. For the data used, a sensitivity analysis is made, and the variance on the residuals is seen to be the most important parameter. In the discussion section a few hints are given for a new measuring programme and further development of the model.

7. TRACER EXPERIMENTS USING SEAWEED

7.1. Introduction

In earlier chapters we have shown that dispersion models depend to a great extent on the time and length scale considered. The simulation model based on current measurements within a limited area can hardly be used for length scales orders of magnitude larger than this area. Furthermore, the assumptions regarding homogeneity, isotropy, etc. do not allow the model to be used in more complicated areas. Therefore, a model describing the initial mixing of effluents from Barsebäck in the Sound will not give reliable results for all parts of the Danish Straits. As shown in Chapter 3, one solution to this problem is to use tracer techniques. For small distances dye (rhodamin) and short-lived radioisotopes have been used successfully as tracers. Due to the detectability of low concentrations, radioisotopes have also been used as tracers covering very long distances.

Furthermore, radioisotopes often make it possible to utilize double tracers. For the Sellafield study, see Chapter 3, both ^{137}Cs (half-life 30 years) and ^{134}Cs (half-life 2 years) can be used and simultaneous measurements of the two at different locations can give information about transport time. One more isotope released from Sellafield can be measured several thousand kilometers away, viz. ^{99}Tc (half-life $2 \cdot 10^5$ years) which has an unusually high affinity for seaweed (CF $\sim 10^5$). It can therefore be used, even though its concentration in water is very low. Using seaweed in the Barsebäck study at least one isotope, ^{60}Co (half-life 5.3 years), can be used in the construction of an empirical dilution/transport model for the Danish Straits.

The investigation of ^{60}Co in the Danish Straits was done for two purposes: For small distances from Barsebäck the ^{60}Co levels in seaweed can verify some aspects of the ellipse model, and where this model is not applicable due to its restrictive assumptions a picture of the steady state situation can be given. In this chapter the results of two collections of seaweed along the

Danish coastlines are described. Additional investigations to improve the use of seaweed as a bioindicator are also described. Two main problems have been found worthy of a detailed analysis. First, the variation in uptake throughout the year and between species, and secondly, the estimation of the important concentration factor (see below).

7.2. ^{60}Co in the Danish Straits

This topic is described by Boelskifte (1985) where results from a seaweed collection along the Danish coastline in 1982 and 1983

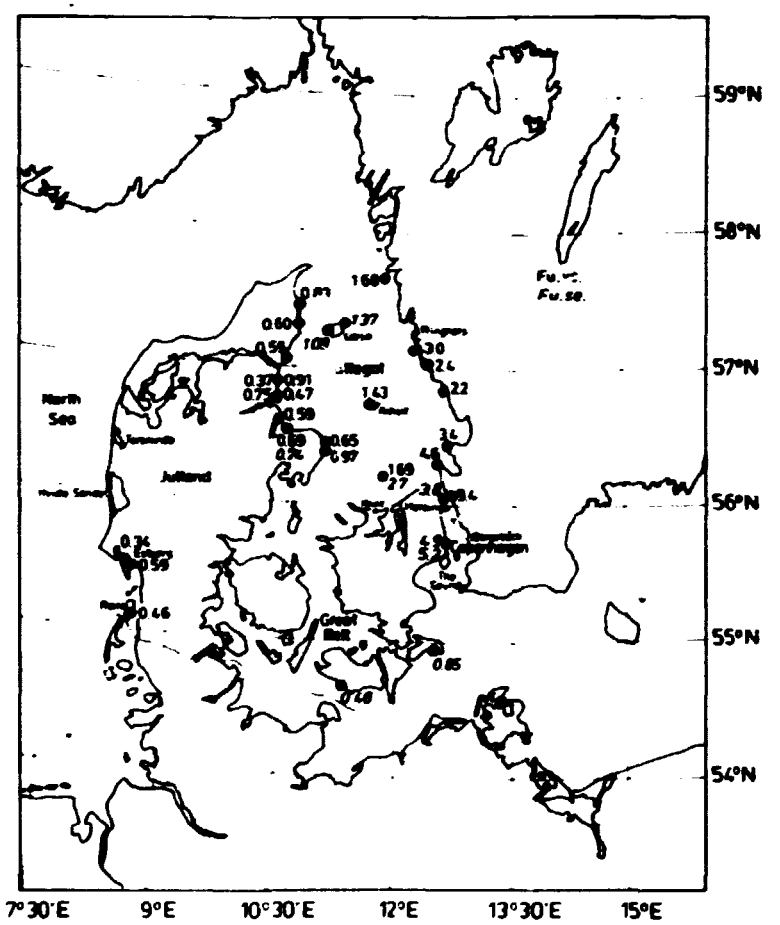


Fig. 7.1. Cobalt-60 in sea plants in 1982. (Unit: Bq kg^{-1} dry weight). Fu. ves.: Fucus vesiculosus. Fu. se.: Fucus serratus.

are detailed. Figures 7.1 and 7.2, taken from the reference, show the similarities between the two years and give a picture of the three areas representing different dilution patterns: the North Sea, Kattegat and the Sound. The initial mixing of ^{60}Co released from Barsebäck takes place in the Sound. The seaweed measurements indicate a narrow plume closely following the Swedish coast. A combination of the results from the Sound and Kattegat leads to an important mathematical description of the level of ^{60}Co in seaweed (dry weight) as a function of the distance in X km, from Barsebäck:

$$\text{Concentration in Fucus } [\text{Bq}^{60}\text{Co kg}^{-1}\text{dry}] = 1200X^{-1.4}$$

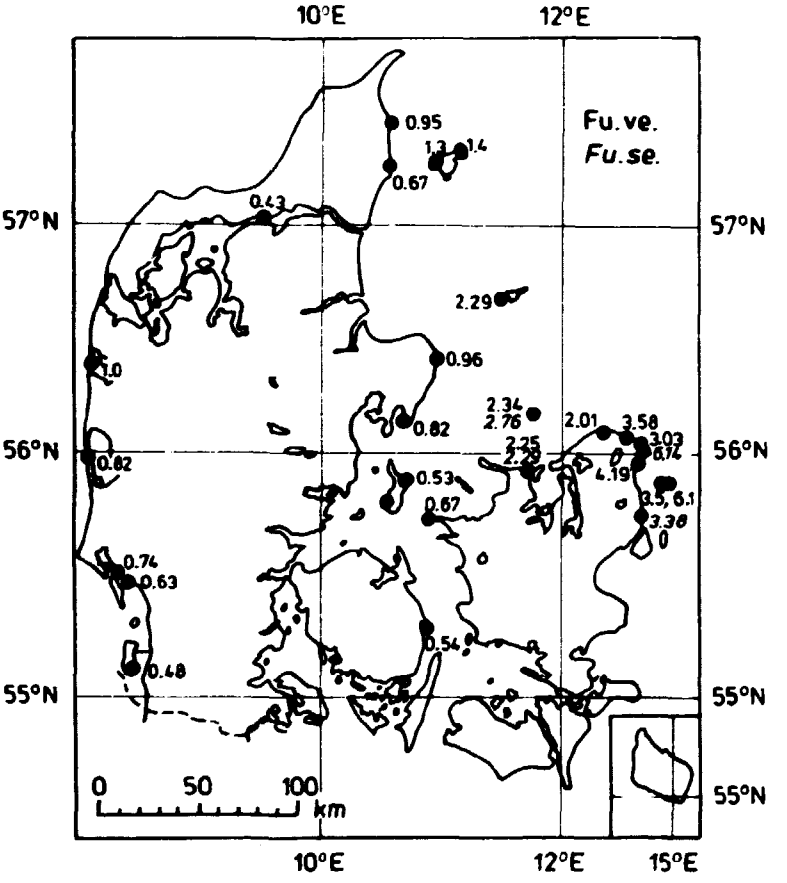


Fig. 7.2. Cobalt-60 in Fucus vesiculosus and Fucus serratus 1983. (Unit: Bq kg^{-1} dry weight) Location names, cf. Fig. 7.1.

This relation is valid whether data from 1982, 1983 or both are used. In the reference it is shown, furthermore, how this relation can be used to estimate the mean residence time in the Kattegat. A figure of two months is found in agreement with oceanographic information.

The ^{60}Co found in seaweed from the North Sea indicates higher levels than expected if Barsebäck or Ringhals were the only sources. Therefore it is discussed how English or French sources might contribute via the North Sea to the levels found on the west coast of Jutland.

Some of the problems using ^{60}Co and seaweed are discussed. ^{60}Co has a high sedimentation rate, but it is argued that since the transport time from Barsebäck to Kattegat is small, only an insignificant amount will be in the sediment. The problems of seasonal variation, differences between Fucus species (F. vesiculosus and F. serratus) and the problem of getting reliable water concentration estimates from the seaweed concentrations are all described in the following sections.

For a continous release of ^{60}Co, dose calculations are easy to make. An example is given in Boelskifte (1984). For a momentaneous release the doses, as described in Boelskifte (1983), depend to a great extent on the current direction at the moment of release. This is due to the position of the Sound as a transition area between the Baltic Sea, with long mean residence time (25 y), and the North Sea, with short mean residence time (2 - 3 y).

7.3. Discussion

7.3.1. Verification of the simulation model

It is surprising that the relationship between Fucus and distance from Barsebäck is valid both for small (< 1000 m) and long distances (> 100 km). The simulation model is only valid for a shorter range (0 - 20 km), but in this area it can be compared

to the seaweed measurements. In Table 6.1 we showed the average of the "area-equivalence" for 100 simulations. Adding a constant mean velocity to the simulated one will move the particles accordingly and thereby we can substitute "time" in the table by "length" and thereby compare results from the two models. We also saw in Chapter 6 that for some sets of initial conditions the size of the cluster would be proportional to $T^{1.3}$. This means that plotting the inverse of the results on a double logarithmic paper will illustrate the concentration in water as function of distance from the discharge, and the slope of the line will be - 1.3. This is very much like the results from seaweed measurements (slope -1.4) and it confirms the reliability of the simulated average dilution. Other kinds of statistics, such as the distance from origo and conditional simulation cannot be verified by seaweed measurements.

This verification raises the question of the reliability of bioindicators. This is the topic of the following three sections, which we find relevant since bioindicators are so widely used in the understanding of transport in the marine environment. Experiments with bioindicators are described for instance in Dahlgaard (1981).

7.3.2. Seasonal variation

In this section we will investigate some of the problems associated with using seaweed as a bioindicator, viz. the variability among samples, the difference between Fucus vesiculosus and Fucus serratus, and the seasonal variation. The variability and the difference between the two species have been tested at a low-level area for ^{60}Co, Klint, (see Fig. 7.1) and near Ringhals power plant. The results are shown in Tables 7.1.a and 7.1.b. The content of ^{60}Co in F. ves. at Klint was 1.27 Bq kg^{-1} (dry) with a standard deviation of 0.14. For F ser. the result was 2.11 Bq kg^{-1} (dry) and S.D. = 0.14, which indicates that the two species differ significantly with respect to ^{60}Co. For ^{137}Cs the difference was also significant,although less pronounced. Also for ^{90}Sr, $^{239,240}Pu$ and ^{241}Am, F.ser. has a higher content than F.ves. But for 99Tc, F.ser. is significantly lower than F.ves. This result

is important because it shows that as a rule, one species cannot be used as substitute for the others in areas where only one is found.

At Ringhals, where the ^{60}Co levels are much higher than at Klint, the samples have been analysed for ^{54}Mn, ^{58}Co, ^{60}Co, ^{65}Zn and ^{137}Cs (Table 7.1b). For ^{54}Mn and ^{137}Cs no difference can be found between the species, but for ^{65}Zn the difference is significant. For ^{58}Co there is no significant difference, and for ^{60}Co further analyses are needed before a conclusion can be made. From these observations, it is obvious that factors related to a specific place can change the loss and uptake of radionuclides significantly. This is true for salinity, temperature, etc. To compare with laboratory experiments see, for instance, Dahlgaard (1983).

The content of radioisotopes in seaweed varies throughout the year. To test whether this seasonal variation can explain some of the discrepancies found, a collection program has been set up where one sample of each species is collected each month, both at Klint and at Ringhals. Due to weather conditions, it has not been possible to get samples from Ringhals for each month. The investigation of seasonal variation has to continue for years before final conclusions can be drawn. But some illustrative results have already been found.

The results are shown in Table 7.2.a and 7.2.b. The concentrations of ^{60}Co in the water at Klint is hardly constant (slow variations must be expected) and not at all constant at Ringhals. Except for a few samples, F.ser. shows higher values than F.ves. In order to produce a reliable picture of the seasonal variation a program must last several years; conclusions cannot be drawn from just one year of analysis. To overcome the problem of varying water concentration, the two species can be used as control for each other.

The ratio of one species to the other for the two locations is shown in Fig. 7.3. This ratio is different between the two places and it varies with time. This can hardly be explained, even though statistical variation can explain a certain amount of the

variation. But the conclusion that each of the two species differs with respect to content of ^{60}Co is obvious.

For ^{137}Cs the figures are different and the water concentrations fluctuate less. At Ringhals the two species are equal except for two months. At Klint the difference is larger but not constant. For both places there is a tendency, at least for F. ves., to show lower levels during the winter than during summer. However, a longer time period is needed to check if that is a significant difference.

Table 7.1a. Radionuclides in the brown algae Fucus vesiculosus (Fu.ve.) and Fucus serratus (Fu.se.) collected at Klint (55°58'N 11°35'E) 24/3 1983 (Unit: Bq kg^{-1} dry weight)

Species	^{40}K*	^{60}Co	^{90}Sr	^{99}Tc	^{137}Cs	$^{239,240}Pu$	^{241}Am
Fu.ve.	24	1.32	5.0	50	8.1	0.29	0.057
Fu.ve.	24	1.26	6.0	151	7.9	0.29	0.035
Fu.ve.	23	1.06	5.0	180	7.6	0.23	0.023
Fu.ve.	25	1.45	5.7	137	8.3	0.32	0.042
Fu.ve.	26	1.28	5.1	187	8.7	0.24	0.030
Mean		1.27	5.4	141	8.14	0.27	0.037
SD		0.141	0.46	55	0.42	0.038	0.013
SE		0.063	0.21	25	0.188	0.017	0.058
Fu.se.	27	1.96	9.0	81	9.7	0.54	0.057
Fu.se.	29	2.1	10.3	87	9.0	0.62	0.046
Fu.se.	32	2.3	8.8	93	11.4	0.50	0.039
Fu.se.	29	2.1	11.1	86	9.9	0.35	0.110
Fu.se.	31	2.1	9.8	79	9.7	0.61	0.073
Mean		2.11	9.8	85	9.93	0.52	0.065
SD		0.137	0.95	5.5	0.87	0.11	0.028
SE		0.061	0.42	2.5	0.39	0.049	0.013

*Unit: g kg^{-1} dry weight

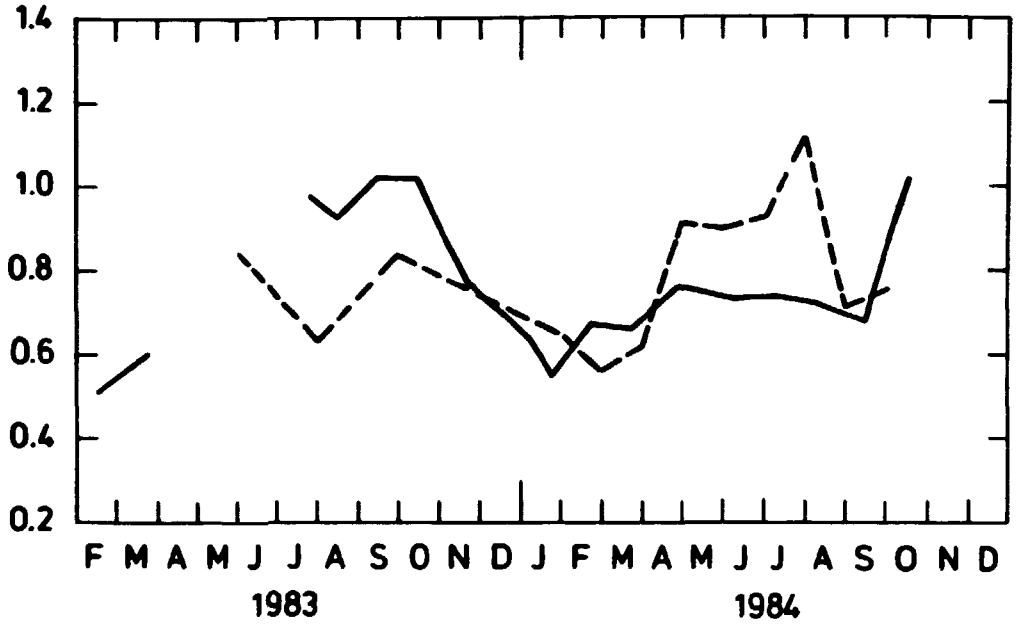


Fig. 7.3. Ratio of ^{60}Co in Fucus versiculosus to Fucus serratus collected at Klint (——) and Ringhals (----).

Table 7.1b. Radionuclides in the brown algae Fucus vesiculosus (Fu.ve.) and Fucus serratus (Fu.se.) collected at Ringhals 2/6 1983 (Unit: Bq kg^{-1} dry weight)

Species	^{54}Mn	^{58}Co	^{60}Co	^{65}Zn	^{137}Cs
Fu.ve.	3.09	9.5	29.9	31.3	10.3
Fu.ve.	2.79	11.1	31.3	34.0	10.2
Fu.ve.	2.92	12.3	38.7	39.3	9.9
Mean	2.93	10.9	33.3	34.9	10.1
SD	0.15	1.43	4.7	4.1	0.22
SE	0.09	0.82	2.7	2.4	0.13
Fu.se.	3.40	11.9	39.4	40.0	11.2
Fu.se.	3.28	12.2	42.2	45.7	8.7
Fu.se.	2.36	10.5	35.8	40.3	10.1
Mean	3.01	11.5	39.1	42.0	10.0
S.D.	0.57	0.90	3.2	3.2	1.28
S.E.	0.33	0.52	1.8	1.8	0.74

Fluctuations in the Cs levels can be moderated by incorporating the ^{40}K levels, while K and Cs are compatible in the metabolism. Tables 7.2a and 7.2b indicate that dividing the ^{137}Cs levels with the ^{40}K levels will eliminate the difference between the species and, especially for Ringhals, make the seasonal variation clearer.

7.3.3. Estimation of a concentration factor for ^{60}Co

The intensive use of bioindicators in environmental studies is based on the assumption that the measured level, C, of an element in the indicator organism in one way or another represents the level in the water, $C_{indicator} = f(C_{water})$, where f usually is assumed to be a constant factor, k. The advantage in using bioindicators is that k is often orders of magnitude higher than unity, i.e. the concentration in the indicator is much higher than in the water, leading to much easier measurements. In case of a steady-state situation k is called the concentration factor, CF, and defined as follows for ^{60}Co:

$$CF = \frac{\text{conc. in seaweed } [Bq\,^{60}Co\ kg^{-1}dry]}{\text{conc. in water } [Bq\,^{60}Co\ \ell^{-1}]}$$

The estimated value of CF is usually based on (several) simultaneous measurements of C_{water} and $C_{indicator}$. This gives reliable estimates as long as a steady state is present, i.e. there is equilibrium between the two concentrations. Otherwise results which are quite misleading can arise.

With this in mind, we decided to set up an experiment to compare an integrated water sample with a bioindicator, Fucus vesiculosus. The experiment concentrated on only one radioisotope, ^{60}Co, since the low level in the water requires a precipitation procedure and thus does not allow for measurements of other elements in the water.

It has been possible to do the experiment simultaneously at two places, the Swedish nuclear power plants Ringhals and Barsebäck (Fig. 7.1 shows the locations). The purpose of the experiment was to

a) Estimate the concentration factor, CF for ^{60}Co between Fucus vesiculosus and water.
b) Examine a method for continous water sampling.
c) Estimate the average concentration of ^{60}Co in water in time periods of one month and compare it with discharge data.
d) Investigate the velocity of uptake for ^{60}Co in uncontaminated seaweed.

The experiment has been carried out at Ringhals since April 1983 and at Barsebäck since December 1983, as shown in Tables 7.3a and 7.3b. At both places the water is sampled from the cooling water channel.

The pump has a capacity of approx. 13 l each 24 hours or about 400 l each month convenient for the analysis. At the end of each month we made a precipitation and measured the precipitate on a Ge(Li)γ-counter. From Barsebäck only one Fucus sample was available, while no natural population was found in the channel and it was necessary to move some plants from a lowlevel area (Limhamn) to Barsebäck. We have collected several Fucus samples from Ringhals each month: one sample of F. ves. as close as possible to the water sampling point, one sample of F. ves. and one of F. ser. at a point on the coast close to the channel (described in Section 7.3.2), and one sample of a transplanted population, moved from Varberg to the cooling water channel. Four moves from Varberg have been made: 1/6 and 1/12 1983, 1/5 and 1/9 1984. During some of the months samples "retransplanted" from Ringhals to Varberg were also measured. The first moved population was moved back to Varberg 1/5 1984.

One of the reason for choosing one month as the integrating time is that the discharges from the plants, which should be compared to the water measurements, are reported monthly.

7.3.4. Results

Tables 7.3a and 7.3b show all the results so far from Ringhals and Barsebäck, respectively. In the following the relevant comparisons are made. First the results from Ringhals are investigated.

Table 7.2a. Seaweed collected at Klint [Bq kg^{-1}(dry)]

Date		F.ves. ^{60}Co	F.ves. ^{137}Cs	F.ves. ^{40}K*	F.ves. $^{137}Cs/K$	F.ser ^{60}Co	F.ser ^{137}Cs	F.ser ^{40}K*	F.ser $^{137}Cs/K$	^{60}Co Ves/Ser	^{137}Cs Ves/ser
14/2	83	1.32	7.47	21.4	0.35	2.59	9.53	29.5	0.32	0.51	0.78
24/3	83	1.27	8.14	24.5	0.33	2.11	9.93	29.6	0.34	0.60	0.82
25/7	83	2.25	9.75	27.2	0.36	2.29	9.91	25.7	0.39	0.98	0.98
18/8	83	2.80	8.13	32.5	0.25	3.01	7.60	26.3	0.29	0.93	1.07
16/9	83	2.66	10.99	39.3	0.28	2.61	6.86	19.4	0.35	1.02	1.60
21/10	83	3.04	7.69	29.8	0.26	2.97	6.57	27.3	0.24	1.02	1.17
24/11	83	2.81	7.17	30.6	0.23	3.67	8.57	33.6	0.26	0.77	0.84
6/1	84	2.53	7.00	27.4	0.26	3.98	9.47	29.9	0.32	0.64	0.74
23/1	84	2.15	7.38	26.4	0.28	3.91	8.47	32.3	0.26	0.55	0.87
24/2	84	2.07	7.17	24.0	0.30	3.07	8.85	30.0	0.30	0.67	0.81
23/3	84	1.90	7.26	23.8	0.30	2.86	8.34	27.2	0.31	0.66	0.87
30/4	84	1.57	8.52	28.8	0.30	2.06	7.11	24.1	0.29	0.76	1.20
8/6	84	1.77	8.58	30.3	0.28	2.41	9.48	30.5	0.31	0.73	0.90
6/7	84	2.16	9.09	29.0	0.31	2.93	9.78	28.8	0.34	0.74	0.93
8/8	84	2.97	9.54	29.3	0.33	4.12	9.52	34.3	0.28	0.72	1.00
17/9	84	2.01	9.14	33.4	0.27	2.96	8.65	26.8	0.32	0.68	1.06
18/10	84	3.51	8.49	33.1	0.26	3.45	9.05	33.4	0.27	1.02	0.94

*Unit: gk kg-1 dry weight.

Table 7.2b. Seaweed collected at Ringhals [Bq kg^{-1}(dry)]

Date		F.ves. ^{60}Co	F.ves. ^{137}Cs	F.ves. ^{40}K*	F.ves. $^{137}Cs/K$	F.ser ^{60}Co	F.ser ^{137}Cs	F.ser ^{40}K*	F.ser $^{137}Cs/K$	^{60}Co Ves/Ser	^{137}Cs Ves/ser
28/3	83	-	-	-	-	17.9	9.98	33.4	0.30	-	-
2/5	83	22.8	9.53	27.3	0.35	-	-	-	-	-	-
2/6	83	33.3	10.13	29.4	0.34	39.1	9.99	28.5	0.35	0.85	1.01
30/6	83	63.1	10.66	29.3	0.36	-	-	-	-	-	-
2/8	83	86.8	9.38	23.1	0.41	137.1	9.34	25.4	0.37	0.63	1.00
1/9	83	68.5	7.57	25.9	0.29	-	-	-	-	-	-
3/10	83	71.8	6.25	29.4	0.21	85.1	5.96	22.4	0.27	0.84	1.05
2/12	83	78.6	7.84	35.0	0.22	106.2	7.60	36.7	0.21	0.74	1.03
1/2	84	67.8	6.64	28.4	0.23	104.7	8.36	38.9	0.21	0.65	0.79
1/3	84	51.3	5.54	19.5	0.28	92.2	8.53	36.1	0.24	0.56	0.65
2/4	84	53.6	6.81	23.1	0.30	86.6	6.62	24.4	0.27	0.62	1.03
2/5	84	39.0	8.17	28.0	0.29	42.7	8.44	28.7	0.29	0.91	0.97
1/6	84	41.9	9.06	28.5	0.32	46.3	9.09	27.2	0.33	0.90	1.00
3/7	84	46.1	9.63	26.3	0.37	49.5	9.59	27.5	0.35	0.93	1.00
1/8	84	41.6	7.06	22.8	0.31	37.0	7.68	25.0	0.31	1.12	0.92
3/9	84	41.2	7.54	29.2	0.26	57.6	8.42	25.6	0.33	0.71	0.90
2/10	84	47.8	9.47	28.8	0.33	63.8	8.74	25.3	0.35	0.75	1.08

*Unit: gK kg-1 dry weight.

7.3.4.1. Fucus/moved Fucus (Fig. 7.4). The first population of Fucus was moved 1/6 1983. About 20 plants from Varberg were marked and placed near the water sampling point at Ringhals. A sample was taken each month, and we found that after three months the plants had reached the level of the local plants. Thereafter, the two populations follow each other closely, confirming the variations from month to month. On 1/4 1984 the rest of the transplanted population was moved back to Varberg to test the excretion of ^{60}Co (Table 7.3a), but due to low water levels during March 1984, the plants nearly died and the results thereafter should be treated with great care.

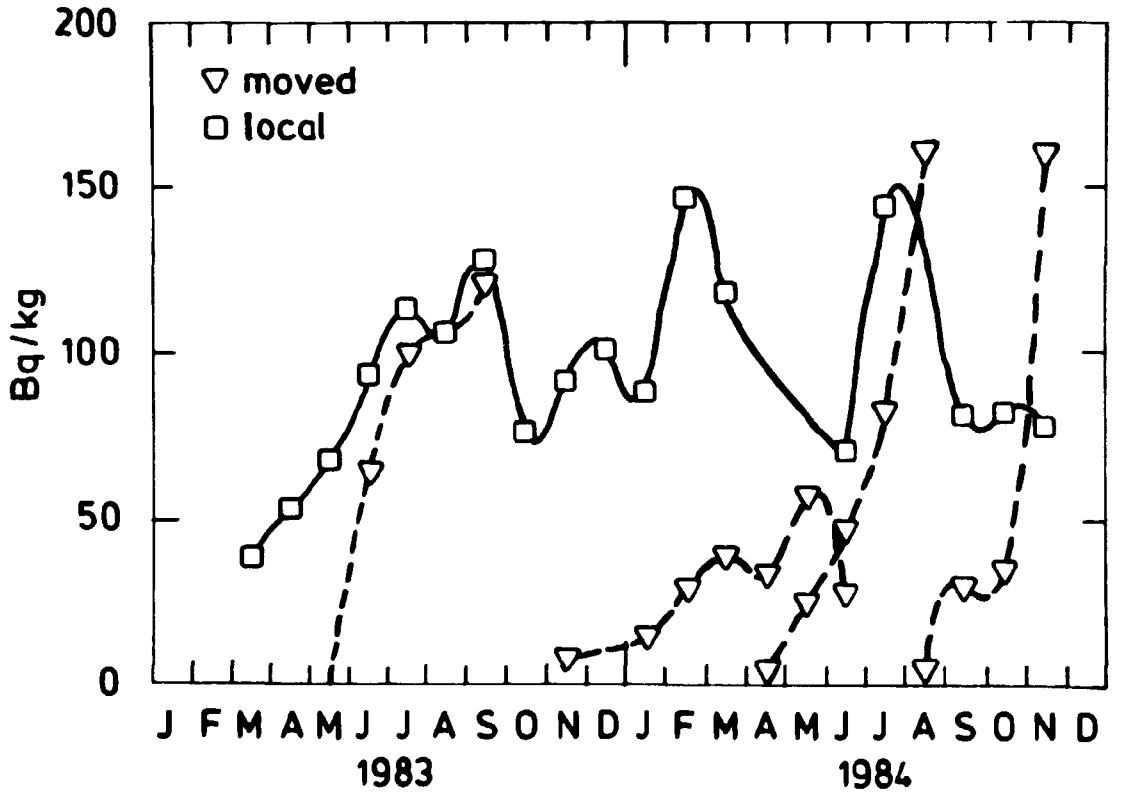


Fig. 7.4. ^{60}Co in Fucus vesiculosus in the cooling water channel at Ringhals. Local population and four transplantations.

A new population also from Varberg was moved on 1/12 1983. In winter the growth rate and the metabolism are small; therefore it takes longer for the plants to reach the local level. In this population there were also some problems with drying, and on 1/7

1984 the last plant was collected. The results for the last three months should be interpreted carefully.

Table 7.3a. ^{60}Co in Fucus, collected water and discharge from Ringhals.

Date	Discharge	Water	Fucus	Fucus moved			
				1983		1984	
	G Bq/ month	Bq/m^3	Bq kg^{-1}	1/6	1/12 Bq kg^{-1}	1/5 dry	1/9
Mar 83	4.22		38.0				
Apr 83	9.01	1.68	53.2				
May 83	4.99	2.20	67.9	1.99			
Jun 83	9.60	2.74	93.5	64.9			
Jul 83	18.0	2.55	114	100			
Aug 83	5.19	0.98	106	106			
Sep 83	9.95	1.66	128	121			
Oct 83	3.1	0.72	76.1	93.4			
Nov 83	4.7	0.62	92.1	105	8.26		
Dec 83	5.3	0.68	102	-	-		
Jan 84	6.5	0.81	88.9	84.9	13.8		
Fec 84	43.2	2.47	147	118	28.6		
Mar 84	2.8	0.34	118	112	395		
Apr 84	3.0	3.08	-		33.6	4.30	
May 84	8.6	3.26	-		57.2	24.8	
Jun 84	5.4	1.08	70.2		27.7	47.0	
Jul 84	20	1.20	144			83.0	
Aug 84	16	1.49	-			161	4.27
Sep 84	4.19	<0.54	81.8			85.0	29.7
Oct 84	4.6	1.12	83.0			101	35.2
Nov 84		<0.15	78.6			106	160
Dec 84		0.62					

The third move, made 1/5 1984, shows the same tendency as the move of the year before: Within three or four months it reaches the local level and thereafter follows the local plants. For other results it is important that, during the growth period, it takes about three months for the plants to reach the actual level. The fourth move was made on 1/9 1984 and it shows a similar tendency.

7.3.4.2. Fucus at two different places. We have investigated the the course of the two F. ves. populations collected in and outside the channel. Each month the samples from the channel show higher levels, but both populations reflect the monthly variations in a similar manner, although the variations are higher in the channel where the levels are also higher.

7.3.4.3. Fucus/water. This relation is the main purpose of the investigation. Both the Fucus and integrated water concentrations are shown in Table 7.3a. If we consider each month separately, it becomes apparent that the concentration factor found in this way is not constant. It differs from approx. $2 \cdot 10^4$ to $12 \cdot 10^4$ (dry weight). The "memory" of the plants must be more than one month, as seen from the transplantation experiment as well. However, a simple mean of the foregoing three months of the water concentration compared to Fucus also gives varying results.

One approach is to incorporate knowledge from laboratory experiments on uptake and loss velocities. In this manner functions can be constructed describing how the concentration in Fucus depends on water concentrations in foregoing months, and how this dependence changes throughout the year. This seems to be a promising method and details can be found in Dahlgaard and Boelskifte (1985).

Due to the presence of algae in the hose there were technical problems with the pumping system during some of the months. These months represent a larger uncertainty than the other

months. It is impossible to make corrections for this lack of water; the only thing to do is to be careful with the interpretation. At least the conclusion can be drawn that when steady state is absent a direct measurement of the concentration factor can at best give only a hint of the order of magnitude.

7.3.4.4. Water/discharge. Differences in the course of the two indicate that monthly differences in the current situation result in a varying transfer from the outlet to the cooling water channel. In many investigations the levels in bioindicators are compared to discharge data. If this is done for short distances from the outlet and short time periods (e.g. a month) it can produce misleading results.

7.3.4.5. Barsebäck. As noted above a similar experiment has been set up at Barsebäck. No seaweed occurs naturally near the water-sampling point and therefore some plants were moved from Limhamn. This move was made on December 1, 1983 but due to technical problems the water sampling first began on January 1, 1984. Two moves have been made in addition to the first, May 1,1984 and September 1, 1984. The quick uptake for the first population is remarkable. The low level found in Fucus in April and thereafter probably is due to low water levels in March, which almost killed the plants. The second move also shows a rapidly increased level of ^{60}Co.

The CF calculated each month shows results between $5 \cdot 10^4$ and $55 \cdot 10^4$, with the former based on Fucus two months after the move. The latter is about five times higher than what is found at Ringhals. This difference between Barsebäck and Ringhals, which has also been seen earlier (Aarkrog et. al., 1982), can hopefully be explained when the experiment has been run for a longer time. Our current hypothesis is that the chemical form of ^{60}Co differs from the two power plants. The difference in salinity also is responsible for at least a part of the different CF-values. The fast accumulation during winter in Barsebäck is also interesting and cannot be explained on the basis of the Ringhals experiences.

7.4. Conclusion

In this chapter we have described some radioecological investigations. If it is assumed that the goal of health physics is to estimate and give tools for reducing doses to man from radioactive isotopes, it is necessary to develop dispersion and transport models covering a span of length scales. In Section 7.2 a model for the Danish Straits is described. A brief discussion of dose calculations is also provided. In the course of the investigation, it became obvious that dose calculations using bioindicators are based on some assumptions which are ripe for further experiments. The seasonal variation in the radionuclides in seaweed is very important. However, in order to determine this variation an experiment lasting at least more than one year would be necessary. For the moment the important result from that experiment is that, with respect to ^{60}Co, Fucus ves. and Fucus ser. are unequal, since F. ser. accumulate more ^{60}Co.

Furthermore, the concentration factor experiment is described in Section 7.3, where levels of ^{60}Co in Fucus are compared to levels found in water collected continuously. Preliminary results show a CF-value (dry weight) for Ringhals between $2 \cdot 10^4$ and $12 \cdot 10^4$, and for Barsebäck between $5 \cdot 10^4$ and $55 \cdot 10^4$. From the literature (e.g. Unscear, 1982) values ranging from 10^4 to 10^5 were found.

Interpretation of the results is difficult, but the method of investigation came up to its expectation. Similarly, the accumulation experiments were interesting and showed reproducable results. Time series analysis have been tried on the data, but to estimate ARMA-models (Chapter 2.2) or detect seasonal variations more data points are needed. A series of 50 values probably would be sufficient.

Table 7.3b. ^{60}Co in Fucus and collected water at Barsebäck

Date	Water	Fucus moved 1983	Fucus moved 1984	
	Bq m^{-3}	1/12	1/5	1/9
			Bq kg^{-1} dry	
Nov 83		207		
Dec 83		348		
Jan 84	1.39	598		
Feb 84	1.20	661		
Mar 84	0.53	158		
Apr 84	-	180	2.23	
May 84	0.54	113	30.02	
Jun 84	2.31		106	
Jul 84	4.49		230	
Aug 84	2.13		193	2.17
Sep 84	7.72		342	255
Oct 84	1.54			194
Nov 84	0.71			140
Dec 84	0.49			106

8. CONCLUSION

The dispersion simulating model presented here is considered as a new tool in estimating doses to the population from releases from nuclear power plants. Referring to the literature it has become evident that most present models for transport and diffusion in the marine environment need a detailed frame of input data. Therefore, if a model could be constructed based on current measurements alone, a new approach to model building would be available.

A current-measuring program has been set up, and the data have been analysed in a number of ways. The distance dependence correlation function for the actual area in the Sound north of Barsebäck was found to be more complex than one might have expected on the basis of the theory, in that it was hard to estimate an overall decreasing function for the correlation. The effect on the correlation of different high-pass filtrations has been examined and the choice of how to separate into a mean current and a turbulence part has been discussed.

The Box-Jenkins method for time series analysis has been applied to the data, and different autoregressive moving-average (ARMA) models have been estimated. Through these analyses the parameters for the simulation model have been procured. The model simulates particle movements using the estimated ARMA-processes as leading factors. The correlation between particles is governed by their mutual distance and the estimated cross correlation function. The model has been used for several types of simulation experiment, and different aspects of the diffusion mechanism have been studied by testing the influence of each parameter involved.

The current-measuring program can serve as an example of what is necessary for a similar investigation at another location, if the proposed improvements are incorporated. It is especially important to obtain a vertical current profile, although often limitations in the number of meters available produces difficulties.

One aspect of the model, the long term average dilution, has been verified by bioindicator measurements (Fucus vesiculosus), and a good agreement is found. From an experiment on the concentration factor between seaweed and water we demonstrated that the variation in this term is hardly understood in detail.

It has been necessary to study three fields of science: hydrology, statistics (time series analysis), and health physics (especially radioecology), and the benefit of the combination of these different areas has been evaluated. Thereby, some suggestions for future work were made, such as a larger measuring program, incor-

poration of vertical dispersion, the use of meteorological data, and an adaptive estimate of the series over a whole year.

ACKNOWLEDGEMENTS

First of all I would like to thank my supervisor, Associate Professor Henrik Spliid for inspiring me and taking interest in my work.

Many people from the Health Physics Department have provided me with invaluable assistance throughout this investigation. I would like to extend a special thanks to Dr. Scient. Asker Aarkrog for his help and guidance and to Henning Dahlgaard for his cooperation in the experiments referred to in Chapter 7.

Furthermore, I would like to express my appreciation to Torben Mikkelsen and Søren Larsen from the Meteorological Department for sharing their insights about turbulence with me.

Finally, I would like to acknowledge the Marine Pollution Laboratory, National Agency of Environmental Protection, for lending us some of their current meters.

REFERENCES

AARKROG, A. (1979). Environmental Studies on Radioecological Sensitivity and Variability with Special Emphasis on the Fallout Nuclides ^{90}Sr and ^{137}Cs. Risø-R-437 (Risø National Laboratory, Dk-4000 Roskilde) 300 p.

AARKROG, A., BØTTER-JENSEN, L., DAHLGAARD, H., HANSEN, H., LIPPERT, J., NIELSEN, S.P. and NILSSON, K. (1982). Environmental radioactivity in Denmark in 1981. Risø-R-469, (Risø National Laboratory, Roskilde) 159 p.

AARKROG, A., BØTTER-JENSEN, L., DAHLGAARD, H., HANSEN, H., LIPPERT, J., NIELSEN, S.P. (1983). Environmental radioactivity in Denmark in 1982. Risø-R-487 (Risø National Laboratory, Roskilde) 145 p.

APPELQUIST, H. and NIELSEN, P.B. (1980). Akkumulering af spildevand i kystnære farvande. (Danish Isotope Centre, Copenhagen) (In Danish).

AUSTIN, H.M., EVANS, D.A. and NORCROSS, B.L. (1984). Time series analysis as a means of examining long term biological and environmental data sets. Int. council for the exploration of the sea (ICES). Statutory meeting, Copenhagen, 8-12 October 1984. Meeting C.N. 1984, paper D1.

BATCHELOR, G.K. (1952). Diffusion in a field of homogeneous turbulence, II. The relative motion of particles. Proc. Comb. Phil. Soc. 48, 345-362.

BENNETT, R.J. (1979). Spatial Time Series: Analysis - Forecasting - Control (Pion Limited, London) 674 p.

BOELSKIFTE, S. (1980). Statistisk udbredelsesmodel for marint miljø. Eksamensprojekt DtH-IMSOR. (Danmarks tekniske Højskole, Lyngby) (in Danish).

BOELSKIFTE, S. (1983). Doses to man from marine products. In: Seminar on the transfer of radioactive materials in the terrestrial environment subsequent to an accidental release to atmosphere. 11-15 April. 1983, Dublin. (Luxembourg) p. 673-83.

BOELSKIFTE, S. (1984). Distribution of ^{60}Co in the Danish Straits as indicated by the brown alga Fucus vesiculosus. In: Seminar on the Behaviour of Radionuclides in Estuaries. 17-21 Sep. 1984, Renesse, The Netherlands. (Commission of the European Communities, Luxembourg) 217-27.

BOELSKIFTE, S. (1985). The application of Fucus vesiculosus as a bioindicator of ^{60}Co concentrations in the Danish Straits. J. Environ. Radioact., 2. 215-227.

BO PEDERSEN, F. (1980). A monograph on turbulent entrainment and friction in two-layer stratified flow. (Inst. of Hydrodynamics and hydraulic engineering. DtH-ISVA-Series paper-25. Technical University of Denmark, Lyngby).

BORK, I. (1977). Model studies of dispersion of pollutants in Lake Vänern. J. [SMHI-RHO-1977-11] (Sveriges meteorologiska och hydrologiska Institut, Norrköping) 13 p.

BOX, G.E.P. and JENKINS, G.M. (1976). Time series analysis: forecasting and control. (Holden-Day, San Francisco) 572 p.

BUCH, E. (1980). Turbulent mixing studies in Danish waters. Rep. No. 44. (Institute of Physical. Oceanography, University of Copenhagen) (Copenhagen).

CLARK, M.I., GRIMWOOD, P.D., COMPLIN, W.C., (1980). A Model to calculate Exposure from Radioactive Discharges into the coastal Waters of Northern Europe. NRPB-R-109. (National Radiological Protection Board, Harwell) 52 p.

CONRADSEN, K. and SPLIID, H. (1981). A seasonal adjustment filter for use in Box Jenkins analyses of seasonal time series. J. Royal Statist. Soc., Ser. C. 30 (2), 172-177.

DAHLGAARD, H. (1981). Bioindicatores for monitoring radioactive pollution of the marine environment. Risø-R-443 (Risø National Laboratory, Roskilde) 134 p.

DAHLGAARD, H. (1983). Transuranics, rare earths and cobalt, zinc and caesium in Fucus vesiculosus (seaweed): Laboratory excercises and field realities. In: International Symposium on the Behaviour of long-lived radionuclides in the marine environment. 28-30 Sep. 1983. La Spezia, Italy. Edited by A. Cigna, C. Myttenaere. EUR-9214. (Office for Official Publications of the European Communities, Luxembourg) 347-355.

DAHLGAARD, H. and BOELSKIFTE, S. (1985). "Sensi": A model describing the accumulation and time-integration of radioactive discharges in bioindicators (Fucus and Mytilus) including seasonal variation. In: Nordisk radioøkologiseminar, 4, 27. February - 1 March, Gol, Norge.

DAHLGAARD, H. AARKROG, A., HALLSTADIUS, L., HOLM, E. and RIOSECO, J. (1984). Radio-caesium transport from the Irish Sea via the North Sea and the Norwegian Coastal Current to East Greenland. In: Symposium on Contaminant Fluxes Through the Coastal Zone. Nantes, France. 14. -16. May.

Danish Hydraulic Institute (1977). System 21: Transport-dispersion model (Hørsholm, Denmark.)

ENGELUND. F. (1969). Dispersion of floating particles in uniform channel flux. J. Hydraul. Division, Am. Soc. Civ. Eng. 95, HY4. 1149-1162.

EVANS, S. (1985). A box model for calculation of collective dose commitment from radioactive waterborne releases to the Baltic Sea. J. Environ. Radioact. 2, 41-57.

FISCHER, H.B., LIST, E.J., Koh, R.C.Y. (1979). Mixing in Inland and coastal waters (Academic Press, New York) 483 p.

FOGARTY, M.J. (1984). Temperature-yield relationships for the Maine American lobster (Homarus americanus) fishery. A time series analysis approach. International council for the Exploration of the Sea. Statutory meeting, Copenhagen, 8-12 October 1984. Meeting C.N. 1984, paper D13.

GRYNING, S.E., (1981). Elevated source SF_6-tracer disperson experiments in the Copenhagen area. Risø-R-446 (Risø National Laboratory, Roskilde) 187 p.

HALLSTADIUS, L., GARCIA-MONTANO, E., NILSSON, U. and BOELSKIFTE, S. (1986). An improved and validated dispersion model for the North Sea and adjacent waters. Submitted for publication in: J. Environ. Radioact.

HARRISON, P.J. and STEVENS, C.F. (1976). Bayesian forecasting. Royal statist. Soc., ser. B. 38, 205-228.

HARVEY, A.C. (1984). A unified view of statistical forecasting procedures. J. Forecasting. 3, 245-275.

HAY, J.S. and PASQUILL, F. (1959). Diffusion from a continuous source in relation to the spectrum and scale of turbulence. In: Atmospheric diffusion and air pollution. Proceedings of a symposium held at Oxford, August 24-29, 1958. Ed. by F.N. Frenkiel and P.A. Sheppard (Advances in Geophysics, 6), (Academic Press, London), 345-365.

HINZE, J.O. (1959). Turbulence. (McGraw-Hill, New York) 586 p.

HOSMER, T.A. (1984). Bivariate time series models of mean monthly temperatures on the Atlantic Continental Shelf. Int. Council for the exploration of the Sea. Statutory meeting, Copenhagen, 8-12 October 1984. Meeting C.N. 1984, paper D6.

ICRP (1977). Recommendations of the International Commission on Radiological Protection. (ICRP Publication, 26). Ann. ICRP 1 (3), 1-53.

Isotopcentralen (1973). Recipientundersøgelser i Fakse Bugt ved Rødvig. (Copenhagen) (In Danish).

JENKINS, G.M. and WATTS, D.G. (1968). Spectral analysis and its applications. (Holden-Day, San Francisco.) 525 p.

KALMAN, R.E. (1960). A new approach to linear filtering and prediction problems. Basic Eng. 82, 34-45.

KALMAN, R.E. and BUCY, R.S. (1961) New results in linear filtering and prediction theory. J. Basic Eng. 83, 95-107.

KOFOED HANSEN, O. and WANDEL, C.F. (1967). On the relation between Eulerian and Lagrangian averages in the statistical theory of turbulence. Risø Report No. 50 (Risø National Laboratory, Roskilde) 52 p.

KOOPMANS, L.H. (1974). The spectral analysis of time series. (Academic Press. New York) 366 p.

KRUSE, H.H., JACOBSEN, T.S., NIELSEN, P.B. (1980). The Belt Project: Physical measurements in the open Danish waters 1974-77, and their storage. (National Agency of Environmental Protection, Copenhagen) 161 p.

KULLENBERG, G. (1974). An experimental and theoretical investigation of the turbulent diffusion in the upper layer of the Sea. Rep. No. 25. (Institute of Physical Oceanography, University of Copenhagen, Copenhagen).

LI, W.K. and McLEOD, A. (1981). Distribution of the residual autocorrelation in multivariate ARIMA time series models. J. Royal Statist. Soc. Ser. B, 43, (2), p. 231-239.

LJUNG, L. and SÖDERSTRÖM, T. (1983). Theory and practice of recursive identification. (MIT Press, Cambridge.)

LUMLEY, J.L. and PANOFSKY, H.A. (1964). The structure of atmospheric turbulence. (Interscience New York). 239 p.

MATTSSON, S., FINCK, R. and NILSSON, M. (1980). Distribution of activation products from Barsebäck nuclear power plant (Sweden) in the marine environment. Temporal and spatial variations as established by seaweed. In: Environ. Pollut. Ser. B. 105-115.

MIKKELSEN, T. (1982). A statistical theory on the turbulent diffusion of Gaussian puffs. Risø-R-475. (Risø National Laboratory, Roskilde) 77 p.

MIKKELSEN, T. LARSEN, S.E. and TROEN, I. (1980). Some puff model ling principles relevant for dispersion calculations in the atmosphere. Risø-M-2258. (Risø National Laboratory, Roskilde) 28 p.

NIELSEN, P.B. (1983). Deterministic and stochastic modelling of dispersion of released effluents in coastal areas. In: 4th Int. Ocean Disposal Symposium, Plymouth, UK, 11-15 April 1983.

NIEMCZYK, S.J., ADAMS, K.G., MURFIN, W.B., RITCHIE, L.T., EPPEL, E.W. and JOHNSON, J.D. (1981). The consequences from liquid pathways after a reactor meltdown accident. NUREG/CR-1596. SAND80-1669 (Sandia Nat. Lab. Albuquerque, New Mexico.)

ODGAARD, J. (1973). Relative diffusion in nonisotropic turbulence. J. Hydraul. Div. Proc. Am. Soc. Civ. Eng., 99, 239-58.

PALMER, M.D., ASCE, A.M. and IZATT, J.B. (1970). Dispersion prediction from current meters. J. Hydraul. Div. Proc. Am. Soc. Civ. Eng., 96, 1667-1679.

PRIESTLY, M.B. (1981). Spectral analysis and time series. Vol. 2: Multivariate series, prediction and control. (Academic Press, London). 2 vols.

RIEPMA, H.W. (1984). Current meter records and the problem of Lagrangean drifts in the Southern Bight of the North Sea. In: Symposium on Contaminant Fluxes Through the Coastal Zone. Nantes, France. 14-16 May.

SLADE, D.H. (ed) (1968). Meteorology and atomic energy. (U.S. Atomic Energy Commission, Oak Ridge, Tenn.) 445 p.

SMHI. Sveriges Meteorologiska och Hydrauliska Institut (1977). Oceanografiska kontrolundersökningar utanför Barsebäcks kärnkraftstation 1976. (Norrköping) 81 p.

SPLIID, H. (1980). MARIMA. Estimation of multivariate time series and transfer function models. DtH-IMSOR-RR-2 (1980) (Danmarks tekniske Hjøskole, Lyngby).

SPLIID, H. (1983). A fast estimation method for the vector autoregressive moving average model with exogenous variables. J. of Am. Statist. Ass. 78, (384). 843-849.

SPLIID, H., JENSEN, S.M. and PEDERSEN, S.B. (1981). Empirical models for the spreading of spills in marine environments. DtH-IMSOR-RR-13 (1981). In: Proceedings from the 4th International time series meeting, ed. O.D. Andersen, (North-Holland, New York).

TAYLOR, G.I. (1921). Diffusion by continuous movements. Proc. Lond. math. Soc. 2, 196-211.

TENNEKES, H. and LUMLEY, J.L. (1972). A first course in turbulence. (MIT Press, Cambridge, Mass.) 300 p.

UNSCEAR (United Nations Scientific Committee on the Effects of Atomic Radiation) (1982). Ionizing Radiation: Sources and Biological Effects, 1982 Report.

VOIPIO, A. (ed.) (1981). The Baltic Sea. (Elsevier Oceanographic Series, 30). (Elsevier, Amsterdam) 418 p.

ÅSTRÖM, K.J. (1970). Introduction to stochastic control theory (Academic Press, New York) 209 p.

Risø National Laboratory **Risø - M - 2566**

Title and author(s)

Dispersion and Current Measurements
An investigation based on time series analysis and turbulence models

Søren Boelskifte

Date April 1986

Department or group
Health Physics

Groups own registration number(s)

Project/contract no.

Pages 151 **Tables** 11 **Illustrations** 41 **References** 64

ISBN 87-550-1216-7

Abstract (Max. 2000 char.)

A model for the simulation of particle movements in water should incorporate the mutual distance dependent correlation. As long as reliable data are accessible a model can be created of the dispersion in a given area from a statistical description of turbulence. Current measurements have been performed in an area north of the Swedish nuclear power plant Barsebäck, and statistical time series analysis have made it possible to estimate multivariate autoregressive moving-average (ARMA) models for these data using the Box-Jenkins method. The correlation structure for the area has been investigated in detail.

Transport and dispersion models for the marine environment are used in estimating doses to the population from the aquatic food chain. Some of these models are described with special emphasis on the time and length scales they cover.

Furthermore, to illustrate the background of the simulation model, short introductions are given to health physics, time series analysis, and turbulence theory.

Analysis of the simulation model shows the relative importance of the different paramaters. The model can be expanded to conditional simulation, where the current measurements are used directly to simulate the movement of one of the particles.

Results from the model are also compared to results from a collec-

Descriptors - INIS:

BARSEBAECK-1 REACTOR; BARSEBAECK-2 REACTOR; BIOLOGICAL INDICATORS; COMPARATIVE EVALUATIONS; COMPUTERIZED SIMULATION; CORRELATED-PARTICLE MODELS; CORRELATION FUNCTIONS; DIFFUSION; DISPERSION; FORECASTING; RADIOACTIVE EFFLUENTS; RADIOACTIVITY; RADIOECOLOGICAL CONCENTRATION; SEAWATER; STATISTICAL MODELS; TIME-SERIES ANALYSIS; TURBULENCE; WATER CURRENTS

Available on request from Risø Library, Risø National Laboratory, (Risø Bibliotek, forsøgsanlæg Risø), P.O. Box 49, DK-4000 Roskilde, Denmark.
Telephone (02) 37 12 12, ext. 2262. Telex: 43116, Telefax: (02) 36 06 09

tion of bioindicators (Fucus vesiculosus) along the Danish coast. The reliability of bioindicators in this kind of experiment is discussed.